Testing the hypothesis that a cytogenetic rearrangement confers susceptibility to Schizophrenia and related disorders

Julie Wilson-Annan

Presented for the degree of Ph.D.
University of Edinburgh
1997

Declaration

I declare that

a) this thesis has been composed by myself

b) that the work is my own, except where otherwise stated

Julie Wilson-Annan

October 1997

Acknowledgements

I am indebted to a large number of people who have been instrumental in the production of this thesis. Firstly, my supervisors Professor David Porteous and Dr Karen Chapman who I would like to thank wholeheartedly for their support and encouragement throughout the course of this work. Secondly, my co-workers in the psychiatric genetics group, particularly Rebecca Devon, Kirsty Millar and Kathy Evans, not only for their endless patience and advice, but for also kindly reading and commenting on much of this thesis.

Great thanks must also be accredited to Lorna Mitchell for her many hours of toil in establishing the Reference Manager database which proved to be a formidable task.

I would also like to thank Ann Doherty and Heather Davidson to whom I owe much of my sanity after this undertaking and Euan Slorach for much valuable discussion and occasional light entertainment!

I would also like to acknowledge the remainder of the "West Wingers" who have lent me great support throughout the course of this work and for which I am truly grateful.

Lastly, I would like to thank my husband Andrew for his endurance, support companionship and ardour without which I would be lost.

Abstract

Schizophrenia is a serious and debilitating mental illness with a lifetime risk of ~1% in the human population. Family, twin and adoption studies have provided strong evidence for there being a genetic component involved in schizophrenia but the nature of this component is still unclear. Genetic linkage studies have failed to provided strong candidate genes involved in schizophrenia. The identification of cytogenetic abnormalities which are associated with disease provides a valuable alternative method for discovering susceptibility genes involved in disease. We have identified a balanced translocation t(1:11)(q43:q14.3) which co-segregates with psychosis in a large Scottish pedigree. A positional cloning strategy has been undertaken to elucidate the relationship between the translocation and the psychiatric diagnosis in this family, based on the hypothesis that a gene or genes involved in mental illness reside at or near the translocation breakpoint on chromosome 1 and/or 11.This thesis takes several approaches to identifying genes residing around the translocation breakpoint which may be involved with the psychiatric diagnosis in this family.

- **Candidate gene approach:**

Two actinin genes ACTN 2 and ACTN 3, were considered as potential candidate genes, having previously been described as mapping to the general region of the translocation breakpoint on chromosome 1 and 11 respectively. High resolution mapping employing somatic cell hybrid and YAC DNA from the region indicated that these two genes were >2Mb from the breakpoint, lying outwith the YAC contigs, and were therefore not considered further as candidates.

- **High resolution mapping of expressed sequence tagged sites from chromosome 11**

This approach utilised resources available from the Human Genome Project to identify transcripts from the chromosome 11 breakpoint region. High resolution mapping of several expressed sequence tagged sites (EST's) which had been described as mapping in the vicinity of the breakpoint on chromosome 11, was achieved using the same approach as with the actinin genes. Of the 18 EST's mapped none were close enough to the chromosome 11 breakpoint to be considered further as candidates.

- **Allelic Association Study using markers from chromosome 1 and 11 breakpoint regions**

An allelic association study was undertaken using two polymorphic markers, identified as lying in close proximity to the chromosome 1 and 11 breakpoint, to test for linkage disequilibrium with a postulated schizophrenia susceptibility gene in a random population of schizophrenia subjects and unipolar depressed subjects compared with a matched normal control population. This study showed no differences in allele frequencies between the affected and control group, evidence against a nearby gene of major effect in the populations studied.

- **Coincident Sequence Cloning**

In order to identify foetal brain expressed sequences from the region spanning the chromosome 1 breakpoint a cDNA selection based technique, coincident sequence cloning was employed. Two products from the library produced were identified as novel cDNA transcripts and were characterised by sequence and expression studies. These can now be followed up as potential susceptibility factors for psychosis in this family.

The analysis of the actinin genes and high resolution mapping of EST's in the region of the translocation breakpoint has allowed elimination of these candidates from being involved in the psychiatric diagnosis in this family. The coincident sequence cloning experiment in contrast has supplied two novel gene fragments whose candidacy can be further evaluated by means of expression studies, sequence homologies and possibly transgenic studies.

ABBREVIATIONS USED

µCi	microcuries
µM	micromolar
µl	microlitres
5HT	serotonin
A	adenine
abs	absorbance
ACh	acetyl choline
AP-SA	alkaline phosphatase-streptavidin
APP	amyloid precursor protein
BAC	bacterial artificial chromosome
BCIP	5-bromo-4-chloro-3-indolylphosphate
bp	base pair
BSA	bovine serum albumin
C	cytosine
cAMP	cyclic adenosine monophosphate
$CaPO_4$	calcium phosphate
CCK	cholecystokinin
cDNA	complementary deoxyribonucleic acid
CEPH	Centre d'Étude du Polymorphisme Humain
cfu	colony forming unit
CIP	calf intestinal phosphatase
cM	centimorgan
CSC	coincident sequence cloning
CsCl	caesium chloride
CSF	cerebrospinal fluid
CT	computed tomography
D1-5	dopamine receptor 1-5
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dH_20	distilled water
DMF	dimethylformamide
DMSO	dimethyl sulphoxide

DNA	deoxyribonucleic acid
DSM	American Diagnostic and Statistical Manual
DTT	dithiothreitol
dTTP	deoxythymine triphosphate
DZ	dizygotic
E.Coli	Escherischia Coli
ECT	electroconvulsive therapy
EDTA	ethylenediamine tetra-acetic acid disodium salt
EE	expressed emotion
EFA	essential fatty acid
EL	end ligation
ERP	event related potential
EST	expressed sequence tag
EtBr	ethidium bromide
FAS	foetal alcohol syndrome
FISH	fluorescence *in-situ* hybridisation
G	guanosine
G6PD	glucose-6-phosphate dehydrogenase
GABA	gamma amino butyric acid
HCl	hydrochloric acid
HF	hybrid fishing
HGMP	human genome mapping project
HLA	human lymphocyte antigen
hnRNA	heteronuclear ribonucleic acid
HOT	hydroxylamine osmium tetroxide
HTF	HpaII tiny fragments
HVA	homovanillic acid
IBD	identity by descent
ICD	International Classification of Diseases
IDDM	insulin-dependent diabetes mellitus
IPTG	isopropylthio-β-D-galactoside
IQ	intelligence quotient
IRD	inter resource duplex
K_2HPO_4	dipotassium hydrogen phosphate
KAc	potassium acetate

kb	kilobase
KH_2PO_4	potassium dihydrogen phosphate
kV	kilovolts
l	litre
LMP	low melting point
LOD	logarithm of odds
LSD	lysergic acid diethylamide
m	milli
M	molar
mA	milliamps
MAO	monoamine oxidase
MAP	microtubule-associated protein
Mb	megabase
MFP	multifactorial polygenic
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MPA	minor physical abnormality
MRC	Medical Research Council
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
MRS	magnetic resonance spectroscopy
MZ	monozygotic
n	nano
Na_3Ci	sodium citrate
NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium hydroxide
NBT	nitroblue tetrazolium
NCAM	neural cell adhesion molecule
NH_4SO_4	ammonium sulphate
NMDA	N-methyl-D-aspartate
OCA	oculocutaneous albinism
ORF	open reading frame
p	pico
PAC	P1 artificial chromosome

PBGD	porphobilonogen deaminase
pBS	Bluescribe plasmid vector
PCP	phencyclidine
PCR	polymerase chain reaction
PEA	phenylethylamine
PEG	polyethylene glycol
PET	positron emission tomography
PFGE	pulsed field gel electrophoresis
PIC	polymorphism information content
PKU	phenylketonuria
PMSF	phenylmethylsulphonyl fluoride
PNK	polynucleotide kinase
PVP	polyvinylpyrolidone
QTL	quantitative trait locus/loci
RDC	Research Diagnostic Criteria
RED	repeat expansion detection
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcriptase polymerase chain reaction
SDS	sodium dodecyl sulphate
SML	single major locus
SSCP	single strand conformation polymorphism
SSR	simple sequence repeat
STS	sequence tagged site
T	thymine
TCA	trichloroacetic acid
TE	10mM Tris.HCl (pH 8 unless otherwise stated), 1mM EDTA
TEMED	N,N,N,N'-tetramethylethylenediamine
TH	tyrosine hydroxylase
THG	Total human genomic DNA
T_m	melting temperature
TYR	tyrosinase
UTR	untranslated region

V	volts
VNTR	variable number of tandem repeats
WT1	Wilm's Tumour 1 gene
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
YAC	yeast artificial chromosome
$ZnCl_2$	zinc chloride

LIST OF TABLES

Chapter Five

Chapter Six

LIST OF FIGURES

Chapter Three

Chapter Four

Chapter Six

TABLE OF CONTENTS

Chapter One

Introduction

1 Introduction

1.1 Preface

Schizophrenia is a serious and debilitating disorder representing one of the severest forms of the psychoses. It is characterised by mental anguish and impaired judgement and is sometimes so debilitating that the sufferer may permanently lose touch with reality. Schizophrenia is found in all countries, societies and cultures and affects approximately 1% of the world population. It is a distressing disorder for both the sufferer and for their family who often have to shoulder the responsibility providing care. Current treatments merely ameliorate the severest symptoms working more efficiently in some sufferers than others and leaving many facing a future of increasing social isolation.

In Britain, mental illness accounts for 14% of the inpatient NHS budget, 23% of the NHS drug budget and approximately one quarter of the homeless population are estimated to have had contact with psychiatric services (MRC field review: Biological Psychiatry 1993). Mental illness therefore poses a significant financial and social problem.

The aetiology of most mental illnesses is complex involving the interaction of various risk factors including genetic predisposition along with social and environmental factors. Although it has been established that a genetic component exists for schizophrenia the nature of this component remains elusive, primarily due to issues such as unknown mode of inheritance, phenotypic uncertainty, genetic heterogeneity, late onset and variable penetrance.

Schizophrenia is both a difficult and challenging disorder to study genetically but recently the field has advanced with new methodologies finally coming to fruition. There is now a real possibility that the genes responsible for the condition can be identified allowing the provision of better therapeutic treatments and understanding for those who are afflicted.

1.2 Schizophrenia

Schizophrenia is the term used to describe one of the principle forms of major mental illness which affects the human population. The term schizophrenia (the splitting of personality) used in 1950 by Eugen Bleuler to denote the breakdown of integration between emotions thought and actions, owes its origins as a separate

clinical entity to Emil Kraeplin, who in the nineteenth century, based on symptoms and long term outcome, divided "madness" (not due to coarse brain disease) into manic depressive insanity and dementia praecox. The latter is known today as schizophrenia.

1.2.1 Symptoms of schizophrenia

The core symptoms of schizophrenia were first described at the beginning of the 19th century by Pinel and Haslam although it is likely that schizophrenia has existed for considerably longer than this. The symptoms were later classified by Kraeplin (1919) into dementia praecox. The symptoms of schizophrenia are highly variable both between individuals and across time in the same individual, often being similar in nature, but very different in form. Some of the most prolific symptoms characterising schizophrenia are listed below. Some individuals will suffer several of these symptoms whereas others will not show clear signs of any, although this may change with the course of their illness, reflecting the syndromic nature of the disorder;

- *Thought disorder.*

  The form and flow of thought is disjointed and the train of thought of an afflicted individual may be difficult to follow due to lack of structure, coherence and a loosening of associations. The thought process may be reduced or slowed and may result in poverty of speech.

- *Abnormal beliefs and delusions.*

  These may take a variety of forms. The individual may feel that their thoughts are being controlled by some external influence or may suffer delusions of persecution or grandeur.

- *Prominent or auditory hallucinations.*

  These are perhaps the most dramatic and certainly the most visible symptoms seen in schizophrenia. Again they can vary widely in terms of their nature and it is believed that the fundamental defect may concern incorrect interpretation of events.

- *Mood disorders.*

  These may interchange between depression, euphoria and anxiety. There seems to be disconnection between mood and other functioning such that moods are often expressed inappropriately. This is often referred to as incongruent mood.

- *Motor alterations.*

  Restlessness, hyperactivity or general increased motor functioning may be displayed or alternatively a reduction in motor activity may be seen with long periods of inactivity and immobility. Alteration of motor activity between these two extremes is often termed catatonic schizophrenia.

- *Alteration of social functioning.*

  This may involve withdrawal from social interaction and may be partial or extreme. Alteration may also occur in terms of breaking social conventions.

Schizophrenia symptoms are often loosely categorised as being either positive or negative in nature. Positive symptoms include delusions, hallucinations, thought disorder such as incoherence, tangential movement of thought and loosening of association, and increased motor function and are termed positive as they are considered to represent excess normal function. Negative symptoms of schizophrenia represent deficit to normal function and are defined as being poverty of thought or speech, loss of emotional responsiveness, blunting of emotions, apathy, attention deficit reduced motor function and social withdrawal (Rao and Moller 1994). The latter often imply a more chronic state.

1.2.2 Diagnosis of schizophrenia

Despite the fact that schizophrenia was described as a separate clinical entity over eighty years ago there is still debate about its definition. With the lack of a reliable, unique biological marker to aid the diagnosis of schizophrenia, psychiatrists are dependent entirely on the symptoms seen at clinical presentation for diagnosing schizophrenia and the other major psychoses. Demarcation of the major clinical picture of schizophrenia and other psychoses is not well defined as there appears to be no point of rarity among them, resulting in distinct blurring of the boundaries of

the disease and making diagnosis difficult. Several classification systems have been devised for diagnosing schizophrenia, many reflecting fashionable ideas of the aetiology of the disease and none proving to be any more valid a definition of the disease than the next. Having a large number of differing criteria for diagnosing the psychoses has been problematic causing interpretation and comparison of data to be difficult if not impossible.

One of the first widely used diagnostic criteria (at least in Europe) was that of Schneider and his so called "first rank symptoms". These symptoms included hearing thoughts spoken aloud, hallucinations in the form of a commentary, thought broadcasting, thought withdrawal or insertion and delusional perception. Such symptoms when present in an individual were considered by Schneider to enable the diagnosis of schizophrenia.

Modern operational diagnostic criteria such as the Diagnostic and Statistical Manual of Mental Disorders (DSM IIIR and recently the updated DSM IV, American Psychiatric Association 1994), Research Diagnostic Criteria (RDC Spitzer 1978) and International Classification (ICD 10, World Health Organisation 1992)) attempt to classify schizophrenia and other mental and personality disorders based on a narrower range of symptoms than those of Schneider and often also on the course of disease. Such systems have been widely accepted throughout the world and have provided a common language within psychiatry. These criteria have superseded Schneiders first rank symptoms having been proved to be both reliable and consistent, although still subjective, perhaps remaining so until more is known about the aetiology of the disorder (Farmer et al 1994, Fogelson et al 1991).

Despite periodic revisions of these criteria with the aim of "improving" the definitions, producing slight alterations in the diagnostic categories, there is some indirect evidence relating to the genetic validity of these operational criteria from McGuffin et al (1984).

1.2.3 Subtypes of schizophrenia

Kraeplin recognised three subtypes of schizophrenia based on his experience of affected individuals - hebephrenic, catatonic and paranoid. A fourth type simple schizophrenia was added later by Bleuler. The DSM IIIR recognises five subtypes of schizophrenia (Gallant et al 1990):

Catatonic Type - The essential feature of this subtype is psychomotor disturbance. Diagnostic criteria for this subtype state that the clinical picture should be predominated by any of the following; stupor, negativism, rigidity, excitement or posturing.

Disorganised type - Marked loosening of associations, incoherence, grossly disorganised behaviour and flat or inappropriate affect characterise this subtype. It is associated with extreme social impairment and a chronic unremitting course. Other classifications term this subtype as hebephrenic.

Paranoid type - Preoccupation with one or more systematised delusion or frequent auditory hallucinations related to a single theme and lack of catatonic symptoms are the basis of this subtype.

Undifferentiated type - This is defined as being a type of schizophrenia in which there are prominent delusions, hallucinations, incoherence and disorganised behaviour and where the criteria for catatonic, disorganised and paranoid subtypes are not met.

Residual type - This is diagnosed by the absence of prominent delusions, hallucinations, incoherence and grossly disorganised behaviour but where there is continuing evidence of signs of the illness - emotional blunting, social withdrawal eccentric behaviour and illogical thinking are commonly seen in this subtype.

The ICD 10 classification is very similar to that of DSM IIIR adding one further category, schizophrenic depression to these five (World health Organisation 1992).

It is generally accepted that the tendency for these subtypes to breed true is weak and that many individuals may change subtype as their illness progresses (Kendell 1987).

An alternative subtyping strategy was proposed by Crow (1980). Crow states; "It seems that two syndromes can be distinguished in those diseases currently described as being schizophrenia and each may be associated with a particular pathological process". Crow termed these two syndromes Type I and II, the former being characterised largely by the positive symptoms of schizophrenia, such as delusions, hallucinations and thought disorder. Crow relates these symptoms to an alteration in dopaminergic transmission. The negative symptoms of schizophrenia such as blunting effect, apathy, poverty of thought or speech and loosening of association characterise the Type II syndrome which may be caused by structural changes in the brain rather than being related to the dopamine pathway. Subtyping in this way may allow appropriate therapeutic treatment in that Type II syndrome

may not be reversible if structural changes in the brain have occurred and therefore may be resistant to neuroleptic treatments (Crow 1980).

1.2.4 The Schizophrenia spectrum

It has long been recognised that the relatives of schizophrenic individuals showed an excess of odd and eccentric behaviour although they are not diagnosed as schizophrenic (Battaglia and Torgersen 1996). These individuals have been shown by Kendler et al (1985) to have an increased rate of schizoaffective disorder, paranoid personality disorder and the atypical psychoses but not unipolar disorder, anxiety disorder or alcoholism. The concept of the schizophrenia spectrum originated from the Danish adoption study by Kety et al (1988) who first designated sub-syndromal disorders as being "borderline" or "latent" schizophrenia. The proposed spectrum disorders include schizoaffective disorder, affective illness with mood incongruent delusions, atypical psychosis, delusional disorder, various personality disorders including schizotypal, schizoid, paranoid and borderline personality (Prescott and Gottesman 1993). The spectrum concept implies a continuum of liability for schizophrenia, those with less liability displaying a less severe disease. There is some evidence from family twin and adoption studies that schizotypal and paranoid personality disorders are genetically related to schizophrenia (Baron et al 1985, Torgersen et al 1984 and 1993, Battaglia and Torgersen 1996).

<u>Kraeplin's dichotomy versus the continuum view point;</u>

Kraeplin's notion that schizophrenia and affective disorder were aetiologically distinct diseases has been widely debated in the literature, being challenged by many who believe that schizophrenia and affective disorder merely represent the two extreme ends of a continuum of liability. The term affective disorder covers a group of illnesses in which mood disturbances predominate. There are two main forms of affective disorder, unipolar and bipolar affective disorder (sometimes called manic depression), the latter suffering from episodes of mania as well as depression. There are several methodologically superior family and twin studies which indicate that schizophrenia and affective disorder have little shared liability, with schizophrenia not aggregating in families of affectively ill probands and vice versa (Baron and Gruen 1991, Kendler et al 1985, Kendler and Gruenberg 1984, Baron et al 1985). However, there is also data which refute this, in particular the

description of identical twins by Dalby et al (1986), one which was diagnosed as having mania and the other schizophrenia (there was no obvious environmental evidence to support non-genetic psychosis). McGuffin et al (1982) similarly reported identical triplets two of which had been diagnosed as having schizophrenia and the third with manic depression challenging the Kraeplin dichotomy. It is of course possible (though perhaps unlikely), as pointed out by Dalby et al, that genetic liability to both conditions is present, with expression being moderated by environmental or physiological factors.

In 1933 Kasanin introduced the term schizoaffective psychosis to describe a subset of individuals who seemed to have symptoms of both affective disorder and schizophrenia. The existence of such a disorder which seems to share common attributes of both conditions has been taken by some as evidence for there being a continuum extending from schizophrenia to affective disorder. Since schizoaffective disorder was first described many studies have undertaken to elucidate the nature of this condition and to establish it as a component of either affective disorder, schizophrenia or as a separate heterogeneous group (Lapensee 1992), but none to a satisfactory conclusion. Most studies of relatives of probands with schizoaffective disorder have found an increased risk of both schizophrenia and affective disorders as compared to normal control members (Prescott and Gottesman 1993). Schizoaffective disorder has been subdivided in several studies into mainly schizophrenic and mainly affective subtypes with some distinction to support this division being found in family studies. However, the question as to whether schizoaffective disorder represents a point on a continuum from affective disorder to schizophrenia remains and can not be resolved conclusively until the aetiology of the conditions are better understood.

1.2.5 Course and long term outcome of schizophrenia

Kraeplin's original classification of dementia praecox was one of chronic progressive illness leading to severe disablement in cognitive and social functioning with few individuals showing improvement at follow-up. Kraeplin did however acknowledge that some patients recovered to near normality. One of the most consistent observations of schizophrenic individuals has been the poor outcome compared to individuals with other psychiatric disorders although the course of the disease is highly variable across individuals. Some affected individuals will have one or several episodes returning to near normality whereas others will endure unremitting

deterioration or an intermediate course with gradually increasing disability. In-between these two extremes lies a spectrum of possibilities with repeat acute episodes and varying degrees of disablement (Wyatt et al 1988). Johnstone (1993) reported that a mean of five admissions for relapse occurred in a group of 532 schizophrenia patients followed for 3-13 years after their initial index episode, showing that in many cases schizophrenia is associated with recurrent or persistent problems. It is however noted that not all patients fared badly and that the outcome was variable. Some studies such as that of Harding et al (1987) in which patients from Vermont State Hospital were followed up indicate a favourable outcome for a significant proportion of patients. The patients, who had been ill for a mean of 16 years and hospitalised for a mean of 6 years, were followed up 32 years later and 68% were reported as having no signs or symptoms of schizophrenia. This study has been criticised for bias in sample selection and it is suggested that this group of individuals were not representative of schizophrenia in general. This study does still imply a good prognosis for at least a subsection of affected individuals.

Hegarty et al (1994) did a meta-analysis of a large number of studies on outcome in schizophrenia completed during 1895 to 1992 and reported that overall less than half of patients diagnosed with schizophrenia show substantial clinical improvement after approximately 6 years of follow-up. They also note that the proportion of patients improving after the mid-century was significantly increased, probably reflecting improved treatments, alteration of diagnostic categorisation and selection bias related to changes in health care. All these factors plus the variable definition of what constitutes a good outcome conspire to make interpretation of data produced on long term outcome of schizophrenia difficult. Such factors must be borne in mind when comparing studies on the outcome of schizophrenia.

Many studies have been done to elucidate which factors influence the course of schizophrenia. One of the most important factors seems to be the level of social competence prior to the onset of illness. Social competence includes such things as getting a job and marriage. Poor premorbid functioning seems to predict a poor outcome (Stoffelmayr 1983). There is some indication that the speed of onset may be a prognostic factor, with insidious onset being related to poorer long term outcome.

A World Health Organisation 2 year follow-up study which looked at determinants of outcome of severe mental disorders (Sartorius et al 1986) reported that the

outcome of schizophrenia differed in developing and developed countries with those in the former faring better. This was confirmed in a later 5 year follow-up study of the same cohort (Leff et al 1992). This study also found that a number of other factors, as well as those mentioned above, were predictive of outcome (clinical or social) such as the occurrence of a life event before onset, the acuteness of onset of symptoms and initial social isolation.

There is also evidence that the chance of relapse can be affected by the emotional environment provided by the family of the afflicted individual. High levels of criticism, over involvement or any degree of hostility towards the affected individual constitute so called expressed emotions (Leff 1992). Frequent contact with relatives who display high expressed emotions increases the chances of relapse whereas contact with relatives who have low expressed emotions seems to be protective (Bebbington et al 1994).

1.3 Epidemiology

1.3.1 Incidence

Schizophrenia is found in all countries, societies and cultures and is associated with a life time risk of approximately 1% (dependent on diagnostic criteria)(Hare et al 1987).

In 1986 the World Health Organisations (WHO) completed a study of the incidence of schizophrenia in 10 countries diagnosing schizophrenia based on Present State Examination (PSE) and ICD 9 diagnostic criteria (Sartorius et al 1986). They reported remarkable stability of incidence rates across widely varying geographical areas, providing strong support for the notion that schizophrenia occurs with comparable frequency world wide. It must be noted however that this incidence rate may include acute and reactive substance induced psychoses as the ICD 9 and PSE do not accommodate for duration of illness unlike DSM IIIR which requires six months duration of illness prior to diagnosis.

There is some evidence for pockets of high incidence of schizophrenia such as in Sweden, Ireland and Yugoslavia (Wyatt et al 1988, Eaton 1991). The reasons for this increase is unclear but may be a reflection of differences in methodology and diagnostic practices.

The stability of incidence rates for schizophrenia across varying countries and cultures makes it an unusual disorder in that variation of disease rates are

commonly seen in epidemiology and often provide clues as to the aetiology of the disorder.

Several lines of evidence suggest that schizophrenia is decreasing in incidence in western countries in recent years. Historical fluctuations in demographic profiles, diagnostic practices and criteria for hospital admissions, make it difficult to resolve this issue (Harrison and Mason 1993, Kendell et al 1993, Stromgren 1987, Eaton 1991). It does appear that there has been a relative decrease in the number of severe cases of schizophrenia seen, with the hebephrenic and catatonic subtypes being rare now (Harrison and Mason 1993, Eaton et al 1991). This is likely to be as a result of better treatment strategies and the availability of antipsychotic drugs.

1.3.2 Age of onset

Schizophrenia is predominantly a disease of early adult life with average onset in the second to fourth decades. The difference in the peak age of onset between males and females has been widely reported and seems to hold true despite the use of various methodologies (Faraone et al 1994). Males have a peak age of onset occurring between 18-25 years whereas females have a broader peak of onset at between 26-45 years (Stomgren 1987). Other studies have confirmed this difference between male and female age of onset (Gorwood et al 1995). It has also been shown that the gender difference in age of onset is present across cultures implying a biological rather than a social aetiology (Beratis et al 1994). There has been much speculation as to what this might be and as to whether there is a female specific protective factor causing the difference in age of onset. There is some evidence to suggest that this factor may be oestradiol and that this may enhance the vulnerability threshold for schizophrenia resulting in a later age of onset in females than in males (Riecher Rossler et al 1994).

The age of onset rather than the time of onset appears to be correlated in siblings lending support to the notion that environmental factors contribute little to this phenomenon (DeLisi 1992).

1.3.3 Seasonal variation in birth rate

Many independent studies have shown that there is an excess (10-20%) of winter and spring births in individuals who go on to develop schizophrenia in comparison to the normal population (Bradbury and Miller 1985). Several explanations of this

phenomenon have been offered including (Pulver et al 1992, Pallast et al 1994, Wyatt et al 1988);

- Seasonally varying factor (such as viral infection) occurs during intra-uterine or post-partum period causing developmental abnormality which affects vulnerable individuals and causes schizophrenia in later life. A neurodevelopmental hypothesis for the aetiology of schizophrenia has been widely postulated and will be discussed in later sections.
- Individuals who are vulnerable to schizophrenia because of genetic predisposition have a biological advantage which is protective against seasonal factors which may otherwise cause mortality.

Dassa et al (1996) report that the seasonal effect is more pronounced in females than males and that it is seen primarily in non familial cases of schizophrenia supporting the hypothesis of an environmental hazard. No consistent seasonality effect was reported for any particular schizophrenia subtype but winter births have been reported to be associated with later onset schizophrenia (Dassa et al 1996).

1.3.4 Gender differences

As mentioned previously, women as a group seem to have a later age of onset of schizophrenia than men. This is not the only factor in which there seems to be a gender difference. Females may also have a better outcome than men (Wyatt et al 1988) and appear to respond better to treatment as a general group compared to males (Waddington et al 1993).

Pre-menstrual exacerbation of schizophrenia symptoms have been noted by clinicians as well as increased difficulties in the post-partum and menopausal period. It has also been noted that females fare better during pregnancy, leading to the suggestion that oestrogen may act as a natural neuroleptic.

There may be a neurodevelopmental sex difference with delayed cerebral development in males making them more prone to neurodevelopmental disorders. Obstetric complications are more likely to be seen during the birth of males than females who later develop schizophrenia, although the reason for this remains obscure (Castle and Murray 1991). There has been some suggestion that such obstetric complications may result in an earlier onset and more severe form of schizophrenia although the evidence for this is weak and refuted by several studies (Smith et al 1995). Male schizophrenics have been reported to have lower IQ and

worse premorbid functioning than females who seem to have shorter stays in hospital and a more affective component to their condition than men.

1.3.5 Reproductive fitness

It has been reported that individuals suffering from schizophrenia have a lower fertility rate than that of the general population (Ritsner et al 1992, Haverkamp et al 1982). Some of this effect may be secondary, related to the decreased marriage of afflicted individuals, due to stresses put on relationships as a result of the disorder. The consistent finding that the prevalence of schizophrenia remains stable despite this reduced fertility is puzzling. Several explanations have been offered;

- Carter and Watts (1971) reported increased fertility rates among the relatives of schizophrenics compared to controls and this was also the finding of Erlenmeyer-Kimling and Paradowski in 1966 when they compared the sisters of schizophrenics with women in the general population. The persistence of schizophrenia across generations may indicate that a selective advantage is bestowed on the relatives (possibly genetic carriers) of schizophrenic individuals, the so called "heterozygote effect". Such an affect is displayed by heterozygotes carrying the sickle cell anaemia mutation which is protective against malaria. There are also reports of schizophrenics being less afflicted by certain conditions such as Rheumatoid arthritis, certain cancers and allergies for example (Knight et al 1992, Carter and Watts 1971). Relatively few studies have been done on such a selective advantage being present and studies on individuals with schizophrenia spectrum conditions which may be associated with gene carrier status are required.
- It has been proposed that schizophrenia may be associated with a high rate of mutation which compensates for the reduced fertility (Ritsner et al 1992). However, schizophrenia occurs at a higher frequency than can realistically be explained by rates of genetic mutation in humans (Wyatt et al 1988). Anticipation is the phenomenon of worsening prognosis and/or decreased age of onset in subsequent generations in a family and has been described in several neurological disorders including schizophrenia (see section 1.9.3). Unstable trinucleotide repeats have been demonstrated as underlying several conditions displaying anticipation and such an unstable mutation could provide a plausible mechanism for a high mutation rate in schizophrenia (Bassett et al 1996).

- It is likely that schizophrenia is not a single genetic entity but a heterogeneous collection of genotypes. It may be possible that if several genes are involved that the mutation rate within the normal range could, by occurring at several loci, maintain the observed high prevalence of schizophrenia (Carter and Watts 1971).

There is some evidence of there being a sex difference in reproductive fitness in schizophrenia with fitness being lower in males than females (Bassett et al 1996). Imprinting may be important in psychiatric conditions although to date it has not been demonstrated in schizophrenia.

Imprinting occurs when there is differential expression of genes depending on parental origin. An example of genomic imprinting occurs in Prader-Willi and Angelman syndromes where both result from microdeletions of the long arm of chromosome 15 but if these deletions are paternally derived Prader-Willi syndrome is manifested where as if maternally derived Angelman syndrome results. It is possible that sex difference in reproductive fitness may obscure imprinting effects (Bassett et al 1996).

1.4 Premorbid indicators of schizophrenia

Schizophrenia has been linked with childhood psychological abnormalities, an excess of minor physical abnormalities and various brain morphologies. It may be possible that individuals who later become afflicted with schizophrenia develop abnormalities of various kinds in childhood which allow them to be differentiated as high risk individuals for schizophrenia.

Done et al (1994) and Jones (1994) observed that individuals destined to become schizophrenic may have a number of childhood abnormalities of personality and intellect. Done et al (1994) investigated social adjustment in childhood of individuals who later became afflicted with schizophrenia, in a prospective follow up study based on ratings of social behaviour given by their teachers at the ages of 7 and 11 years. Ratings obtained from individuals followed up in the national child development study who later went on to develop schizophrenia were compared to those obtained from controls from the same cohort. The results indicated that abnormalities of social adjustment were detectable in some individuals who develop schizophrenia in later life ($p<0.01$). Done also reported gender differences in the abnormal social behaviour seen, with girls appearing more withdrawn by the age of

11 years, whereas boys tended to be more hostile and anxious for acceptance than peers of this age.

Jones and colleagues (1994) studied sociodemographic, neurodevelopmental, cognitive and behavioural factors in individuals who later developed schizophrenia, in a British 1946 birth cohort. They found that those individuals who later developed schizophrenia had delayed attainment of speech and problems with speech until early adulthood, they reached motor milestones at a later age than normal, were more socially withdrawn and anxious in adolescence with poorer educational achievement in both childhood and adolescence. This, the authors suggest, indicates that as with some other adult onset disorders the origins of schizophrenia may be found in early life, but they do not exclude a dynamic process over a long period of time.

There have been attempts to associate premorbid childhood behavioural characteristics with abnormalities in brain morphology (Walker et al 1996). Walker et al (1996) claim mild associations with structural brain abnormalities are apparent in children with behavioural problems who later develop schizophrenia. Neurological deviation in new borns at psychiatric risk has been shown by Blennow and McNeil (1991).

1.5 Markers of schizophrenia

Several psychophysiological traits associated with schizophrenia have been proposed as genetic markers for vulnerability to schizophrenia;

Eye Tracking:

Smooth pursuit eye tracking dysfunction has been reported to be present in 51% to 81% of schizophrenia individuals as opposed to 8% of the normal population (Blackwood et al 1991). Normally as an individual follows an object across a visual field a smooth pursuit eye movement system maintains the image at the centre of the retina. Several eye movement systems can be distinguished; smooth pursuit which is comparatively slow, the rapid eye system or saccadic and the vergence system. All three are implicated in the following of a moving target. On following a moving target the smooth pursuit system is initiated but is delayed causing the eye to lag behind the movement. Saccadic eye movements serve to fix the fovea again on the target. When the smooth pursuit system is turned on the saccadic system should be switched off but in a large number of schizophrenics and their relatives

this is not the case causing a high prevalence of both saccadic intrusion and tracking. Impaired eye tracking can be reliably measured using a number of devices and the abnormalities seem to be unaffected by neuroleptic drugs or clinical state and have been shown to be highly heritable and although not specific to schizophrenia seems to be found with a higher incidence than other psychiatric conditions (Blackwood et al 1991, Holzman 1992, Keefe et al 1991). It has been reported that heavy smoking among schizophrenics can normalise smooth pursuit eye movement (see section 1.10).

Event related potentials;

Event related potentials refer to the broad spectrum of brain electrical activity which occurs on sensory, motor or cognitive stimulation and can be measured using an electroencephalographic wave signal. The most pronounced changes reported have been in P300 wave (so called as it represents a positive polarity wave occurring at 300ms after stimulation) amplitude, which is reduced, and P300 wave latency, which is delayed, after exposure to unexpected auditory or visual stimulation (Holzman 1992, Blackwood et al 1991). Family and twin studies have confirmed the heritability of this response (Holzman 1992).

P50 event related potential response

In this test two stimuli are given, the first a "conditioning stimulus" and a second the "test stimulus" (usually auditory). The response to the second stimulus is normally smaller in magnitude to that of the conditioning stimulus. In schizophrenics and their relatives this response is not seen and the response to the second stimulus is similar to that of the first (Holzman et al 1992, Freedman et al 1997). Smoking has been shown to cause a deficit in the inhibition of the P50-evoked response to repeated auditory stimuli in schizophrenic individuals (see section 1.10).

Continuous Performance Test

Continuous performance testing refers to a group of visual vigilance tasks in which a continuous series of stimuli, presented briefly and in close succession, are monitored. The subject is required to respond when a pre-designated target stimulus or sequence appears. Schizophrenic individuals appear to make significantly more errors (failure to respond to target or inappropriate response) and

also have a slower response time than control subjects. This poor scoring is also seen in the offspring and siblings of schizophrenic individuals.

Cognitive deficit markers such as those described above may have a role in improving phenotypic definition and clarifying the boundaries of schizophrenia spectrum disorders as well as identifying gene carriers for linkage analysis and other genetic studies. However, more studies are needed before they can fulfil such roles.

1.6 Neuropathology

Structural abnormalities in the brains of schizophrenics were described as early as 1927 using Pneumoencephalograms and have been described intermittently ever since.

Various brain imaging techniques have become available in the last decade which have allowed *in vivo* brain structure and function to be investigated in patients with schizophrenia and normal controls. Computerized Tomography (CT) and magnetic resonance imaging (MRI) have been used widely to study brain structure, the latter being able to produce three dimensional reconstruction of the brain and excellent grey/white matter resolution. Single Photon Emission Computed Tomography (SPECT) and Positron Emission tomography (PET) have been used to assess functional activity of the brain. Such techniques have improved our knowledge of the pathophysiology of mental illness by demonstrating structural and neurochemical abnormalities in a wide range of mental disorders (Andreasen 1988).

The most consistently reproduced structural brain abnormality seen in schizophrenia is that of ventricular enlargement (Johnstone et al 1976). The ventricles are fluid filled cavities in the brain and their enlargement is assumed to accompany a decrease in neuronal tissue and thus would be expected to indicate a compromise in brain function (Gershon and Rieder 1992). Only 6-40% of schizophrenics show ventricular enlargement (Syvalahti et al 1994) and it is neither specific or consistent with overlap occurring between affected and normal control groups (Suddath et al 1990). Since ventricular volume increases with age it is important that the control group is closely age matched to the affected group (Cleghorn et al 1991). Increased ventricular size appears to be present before the onset of psychosis and is non progressive in nature - as determined by several studies on affected individuals at various stages of illness, raising the possibility that

cerebral abnormalities occur early in development (Heyman et al 1992). The hypothesis that schizophrenia is a neurodevelopmental disorder has been put forward by several groups and has some evidence to support it. This will be further discussed in section 1.10.

Suddath et al (1990) studied monozygotic twins discordant for schizophrenia and found that enlarged ventricles and reduced hippocampi were consistent neuropathological features of schizophrenia. Since monozygotic twins share identical genes many have postulated the underlying neuropathological process can not be wholly genetic. Two points are worth noting here. Firstly, although monozygotic twins share identical genes they may not share identical environments either pre- or post-natally which may result in developmental differences (Gilmore et al 1996). Effects of pre-natal environment has been shown to be of greater importance than previously though in a recent study on the heritability of IQ (Devlin et al 1997). This study showed that prenatal environment could account for as much as 20% of covariance in IQ heritability between twins and 5% between sibling. Studies have shown that the offspring of the unaffected twin still have a higher than normal risk of developing schizophrenia raising the possibility that a responsible gene may not be expressed in certain individuals (reduced penetrance) and that these individuals may act as gene carriers.

Other abnormalities of brain structure reported in schizophrenia include decreased volume of temporal lobe structures such as the medial temporal lobe structures hippocampus, parahippocampus and amygdala (Waddington et al 1993, Murray et al 1992, Falkai et al 1988, Cernansky 1991).

Postmortem studies have indicated that brain weight and length are reduced in schizophrenic individuals compared to controls.

Cellular changes which have been reported include decreased tissue quantity (grey matter) and disorganisation of pyramidal and other neuronal cells in the hippocampus and other limbic circuit structures. The hippocampus develops from the ventricular zone exclusively and is half its adult volume at birth and adult size at approximately two years. The smaller volume and decreased disorganised cell structure are compatible with a developmental abnormality in schizophrenia (see section 1.10).

Jakob and Beckmann (1986) reported that pre-alpha neurones were displaced from their normal position in the upper layers of the entorhinal cortex to the parahippocampal gyrus, suggesting a disturbance of neuronal migration. A large

proportion of neuronal cells usually die in the developing nervous system, a process which seems to eliminate early errors of connection. This process of neuronal pruning has been suggested to be faulty in schizophrenia resulting in the persistence of immature patterns and connections of these cells (Murray et al 1992).

It has been suggested that some structural brain abnormalities may be more profound in the left hemisphere than the right hemisphere (Bracha 1991, Jakob and Beckmann 1986, Crow 1990). These include ventricular enlargement, temporal lobe size reduction and tissue loss. These changes are often subtle. During normal development the left hemisphere lags behind the right hemisphere until the third trimester when the two become equal. If an insult occurs during the "lag" period it is possible that the less developed left hemisphere may be affected to a greater extent than the right.

Several groups have shown gliotic changes, which would be expected to be found in association with cellular abnormalities caused by damage to the central nervous system, are absent in the brains of schizophrenics. The absence of this response seems to suggest a developmental rather than a degenerative process (Murray et al 1992).

Functional changes

SPECT (Single Photon Emission Computed Tomography) has been used to examine the brain at work. Decreased cerebral blood flow in the brains of some schizophrenics has been reported by Ingvar and Franzen (1974). This is known as "hypofrontality". More recently Weinberger's (1987) group used SPECT to look at cerebral blood flow in subjects who were taking the Wisconsin card sort test which is a test of working memory and abstract thinking. In normal subjects this test elicits an increased blood flow in the frontal cortex. In schizophrenics less of an increase in cerebral blood flow is seen and they do less well in the test than normal controls. By identifying various tasks which produce specific patterns of activation in the brains of normal individuals, such as the Wisconsin card sort test, it should be possible to look at specific brain regions which are defective in schizophrenic individuals. Many of these studies have been criticised on the basis of sample selection and lack of attention to medication effects.

1.7 Neurochemistry

Aberrant neurochemistry has long been suspected as having a pathological role in the aetiology of schizophrenia and has lead to a plethora of studies investigating neurotransmitters and their biochemical pathways. Although no neurotransmitter has been directly causally related to the pathology to date many transmitters such a dopamine and serotonin continue to be studied. Further discussion of neurotransmitters can be found in section 1.11.

1.8 Treatment

Antipsychotic drug treatments

Neuroleptic (antipsychotic) drugs were introduced to psychiatric practise in the 1950's and undoubtedly improved the short term prognosis for schizophrenia, apparently altering the course of the disease and preventing further deterioration and relapses.

The neuroleptics are a diverse family of compounds which are all effective in treating some of the symptoms of schizophrenia. They appear to most successful in treating the positive symptoms of schizophrenia, such as delusions and hallucinations although they (in particular the atypical neuroleptic Clozapine) may have some benefit in the management of a subgroup of negative symptoms.

30% of patients taking neuroleptic drugs show only a limited improvement in acute treatment trials but 7% do not respond even to lengthy treatment. Long term maintenance therapy seems to prevent relapses in the majority of patients but 40% will have a psychotic relapse within one year in comparison to two thirds who are not drug maintained (Johnstone 1993).

Many patients experience side effects whilst taking neuroleptics in both long and short term treatment strategies. Side effects include weight gain, lethargy, anticholinergic side effects and in some cases symptoms resembling those seen in Parkinsons disease such as akathisia (motor restlessness) and dystonia (muscle spasm). Parkinsonian like symptoms persist in 20-40% of patients receiving neuroleptic treatment and is a primary reason for non compliance. Tardive dyskinesia (abnormal involuntary movements) are also experienced by a subgroup of patients on long term treatment and there has been a suggestion that this is related to chronic schizophrenia in which there is there has been structural changes in the brain (Strange 1992).

Electroconvulsive therapy (ECT)

Electroconvulsive therapy was first used in the 1930's to treat schizophrenia. It consists of passing an electric current through the brain of a patient in order to produce a convulsion. Control trials conducted by Taylor and Fleminger (1980) showed ECT to be useful in treating schizophrenic patients in conjunction with drug therapies but the effect was short lived, disappearing after 12 weeks.

ECT treatment is now largely confined to treating severe cases of depression in which it appears to have a more beneficial effect.

Non-physical treatments

Non-physical treatments based on psychological principals are advantageous to a small group of patients suffering from schizophrenia. Social and psychological treatments for schizophrenia have been described which aim to both restore the patient to normal social functioning and to remove any precipitators of relapses. Stress is a well known precipitator of schizophrenia episodes. Major life events which cause an individual to have to adapt psychologically can also trigger relapses of schizophrenia. The home emotional environment has been shown to be of importance in relapse and as mentioned earlier, high expressed emotions of relatives is undesirable (Bebbington and Kuipers 1994, Leff 1992, Spring 1981). Interventions capable of modifying the behaviour of family members have proved to be successful in reducing relapses. Psychotherapy of various types has been used to improve the coping skills of afflicted individuals to improve stress management (Cullberg 1991).

A combination of these treatments, both antipsychotic drug and psychotherapy, have proved to date to be the most effective way of treating schizophrenia.

1.9 Aetiology

1.9.1 Genetic Aetiology:

Family, twin and adoption studies have provided strong evidence for there being a genetic component involved in the aetiology of schizophrenia.

Family studies

If schizophrenia aggregates in families it would be expected that the relatives of affected individuals would manifest a higher incidence of the disease and its

associated disorders than that seen in the general population. The risk should also increase as the number of shared genes increases therefore first degree relatives should be at greater risk than more extended generations (e.g. grandchildren) (Tsuang et al 1991).

The majority of family studies done in schizophrenia have shown this relationship to be true. The exceptions, Pope et al (1982) and Abrams and Taylor (1983) which found similar rates of schizophrenia among family members of affected individuals, have been strongly criticised by various authors on methodological grounds.

Although the figures from individual studies vary, depending on particular methodology and diagnostic criteria used, generally the life time risk of schizophrenia in the close relatives of schizophrenic individuals is significantly higher than the general population rate of approximately 1%.

Tsuang et al (1991) assessed the risk of schizophrenia in first degree relatives of schizophrenics to exceed that of the general population by 5-10 times. Prescott and Gottesman (1993) estimates the figure to be 3-4 times and a similar figure is offered by McGue et al (1989).

Several authors have combined data from several family studies to provide consensus average figures for familial risk. The following table is an adaptation of that produced by Gottesman 1991 based on European studies done between 1920 and 1987;

TABLE 1.1: Average Risks for developing schizophrenia amongst relatives of schizophrenics (From Gottesman 1991)

Relation	Lifetime Risk %
First degree relatives	
Parents	6
Siblings	9
Siblings with 1 schizophrenic parent	17
Children	13
Children with 2 schizophrenic parents	46
Monozygotic twins	48
Dizygotic twins	17
Second degree relatives	

Uncles and Aunts	2
Nieces and Nephews	2
Grandchildren	5
Third degree relatives	
First cousins	2
General population	1

There would seem to be little doubt that schizophrenia manifests as a familial disorder but family data alone are inadequate to infer genetic causality as they do not distinguish between genetic and environmental influences.

It is evident that the risks of developing schizophrenia seen in family members do not conform to those implied by classic Mendelian inheritance in which first degree relatives would be expected to have 50% genetic risk and second degree relatives a 25% risk. This implies that a complex genetic and/or environmental pattern of transmission is occurring. Twin and adoption studies have proved useful in determining the extent of genetic versus environmental influence involved.

Twin studies

Monozygotic twins share 100% of their genes whereas dizygotic twins share on average half of their genes. If the assumption that monozygotic and dizygotic twin pairs share a common environment is made, then any difference in resemblance of schizophrenia pathology between the types of twins is attributable to their difference in genetic similarity. Although the principal is simple, carrying out twin studies in schizophrenia is complicated by the relatively low numbers of twins available, both affected and normal, difficulties in phenotype definition and errors in determining zygosity. Despite these complications a number of well designed twin studies have been done.

A twin pair is termed concordant if both members of the pair are affected and discordant if only one member of the pair is affected. The concordance rate is the probability that a co-twin is affected given that his or her co-twin is affected (La Buda at al 1993).

Concordance rates vary widely according to methodology and diagnostic criteria used to determine them but generally despite many different methodologies and diagnostic criteria all showed at least a two fold increase in concordance between

monozygotic twins as compared to dizygotic twins. Among the best methodological studies are the Scandinavian twin studies done by Tienari (1963), Kringlen (1967) and Fischer (1969) based on accurate and well kept national twin and health records. Concordance rates reported in these studies were as follows;

Table 1.2; Twin concordance rates reported by three Scandinavian studies.

	Monozygotic	Dizygotic
Tienari (1963)	15%	7%
Kringlen (1967)	31%	8%
Fischer (1969)	24%	10%

In general the reported concordance rates lie between 35-58% for monozygotic twins and 4-27% for dizygotic twins (Prescott et al 1993, Tsuang et al 1991, Kringlen et al 1993, Carter and Watts 1971). Prescott et al (1993) proposed a median concordance rate based on the results of several studies which had similar methodology, of 48% for monozygotic twins and 15% for dizygotic twins.

It is important to note that no concordance rates of 100% have been reported implying that although there is undoubtedly a substantial genetic component involved there may also be a non genetic factor which may also be important in at least some forms of schizophrenia. Alternatively, reduced penetrance of the genetic component in the unaffected twin may offer a plausible explanation for the reduced concordance rates. Fischer et al (1971) reported that the offspring of the unaffected twin of a discordant monozygotic pair had the same risk of schizophrenia as the offspring of the affected twin which may imply that the unaffected twin is carrying genetic liability to schizophrenia allowing it to be passed onto the offspring but it for some reason is not being expressed (incomplete penetrance). Subtle differences in environment may explain this.

Taken together the twin studies, despite their methodological differences, suggest a major role for genetic factors in the familial aggregation of schizophrenia.

Adoption studies

Adoption studies effectively allow separation of the genetic and environmental components involved in disease since close biological relatives do not share the same postnatal environment (assuming they have been adopted away to non related individuals very early in life). There are several adoption study designs. The prevalence rates of schizophrenia among adoptees born to schizophrenic parents can be compared to those of non schizophrenic parents. One of the earliest studies of this type was that of Heston et al in 1966 which compared a group of adults born to schizophrenic mothers and who were separated two weeks after birth to a control group from non schizophrenic parents (adoptees study). Schizophrenia was diagnosed in 5 of the 47 individuals born to schizophrenic mothers compared to none in the control group of 50 individuals. In addition, 50% of the group born to schizophrenic mothers had some sort of psychosocial disability as adults.

Similar results were reported by Kety et al (1971and 1978) and more recently by Tienari et al (1994) who both compared prevalence of schizophrenia in the adopted and biological relatives of schizophrenic offspring given up for adoption (the adoptees relatives study) and reported an excess of schizophrenia among the biological relatives and not the adopted relatives and non biological control group. Kety's study in 1978 also included parental half siblings of the adopted schizophrenic individuals. Half siblings share approximately 25% of there genes and since these siblings were paternal (i.e. related by a common father) they do not share common prenatal, perinatal or neonatal environments. The finding that these individuals have a higher risk of developing both schizophrenia and related disorders strongly indicates genetic factor in the transmission of schizophrenia.

Adoption studies which have looked at normal children adopted to schizophrenia parents have done much to dispel the belief that schizophrenic like behaviour can be "learned" - so called "vertical cross transmission". Such children show no higher risk of developing schizophrenia than the general population (Tsuang et al 1991).

These results are consistent with shared genes rather than shared environment underlying the increased risk of schizophrenia and associated disorders in the relatives of schizophrenic subjects. Although each of these methods, family, twin and adoption, can be individually criticised in terms of there methodology, together they provide strong accumulative evidence to support the importance of genetic influences on the vulnerability to develop schizophrenia.

1.9.2 Mode of genetic transmission

Having established the importance of the genetic contribution in the causation of schizophrenia researchers have attempted to determine how the schizophrenia gene(s) are transmitted from generation to generation. There is no general agreement on the mode of genetic transmission of schizophrenia and virtually every genetic mechanism has been proposed at one time or another to explain the pattern of inheritance. The classical Mendelian model of inheritance does not adequately explain the genetic transmission seen in schizophrenia. If schizophrenia were caused by a fully penetrant dominant gene it would be expected that 50% of offspring with one schizophrenic parent would become schizophrenic. The observed value is much lower (~16%). If schizophrenia were caused by a fully penetrant recessive gene it would be expected that 100% of children with two schizophrenia parents would be schizophrenic, again this value is far lower (~46%). Therefore it is apparent that more complex models are needed to describe the mode of genetic transmission in schizophrenia.

Three major models have been proposed;

I. Single Major Locus Model:

In its simplest form this model proposes that a single gene with complete penetrance is responsible for the transmission of schizophrenia. A major problem with this model is the explanation of the non Mendelian distribution of the disorder as based on family and twin studies. This can be overcome by assuming reduced penetrance (gene carriers may not necessarily manifest the illness) and allowing for the existence of phenocopies (affected individuals need not have the disorder-associated allele and may be environmentally induced). In more complex models the addition of an environmentally related liability in the form of a threshold construct is introduced. In the single major locus threshold model all individuals are assumed to have a liability or predisposition to developing schizophrenia which is determined by their genotype as well as environmental factors. If liability exceeds the threshold level then the individual will become schizophrenic and if not they will not develop schizophrenia.

There is generally little support for a single major locus model in the literature and several studies have shown this model to be incompatible with the overall rates of schizophrenia found in family and twin studies (O'Rourke et al 1982). Many studies which provide statistical tests for model adequacy reject this model (Faraone 1985,

McGue 1989 Tsuang 1991). It also seems unlikely that a collection of single gene disorders can explain the genetic liability to schizophrenia. However, the rejection of any genetic model may merely indicate that some of the non genetic assumptions of the model are incorrect. For this reason many feel that such a model can not be ruled out entirely (Faraone 1985).

II. Polygenic Models;

Polygenic models assume that genes found at more than one locus are responsible for the familial pattern of schizophrenia. The polygenic model can be further sub-divided into Limited Loci Polygenic (LLP) and Multi Factorial Polygenic (MFP);

a. Limited Loci Polygenic model (LLP)

This model assumes that two or three loci are responsible for the genetic transmission of schizophrenia. The application of quantitative techniques to genetic models involving several genes segregating at several loci is hindered by the fact that there are hundreds of possible loci models. Consequently modelling is done on the basis of the biological significance of the model or on hypotheses formulated about schizophrenia.

b. Multi Factorial Polygenic Model (MFP)

This model assumes that all individuals have a liability or predisposition to developing schizophrenia. If this liability exceeds the threshold level then they will develop schizophrenia. The model assumes the liability is composed of genetic and environmental components each of which have small but additive effects. This model differs from LLP model as it does not specify the number of loci involved, instead it assumes there are many interchangeable loci. Many investigators favour this model as it appears to account for several findings in family studies of schizophrenia not explained adequately by other models, such as the varying degrees of severity and the familial risk seen in schizophrenia. Twin studies are most consistent with this model. One of the draw backs of this model however is its inability to differentiate the relative contributions of genetic and environmental elements. This problem is being tackled at least in part by "path analytical" MFP models developed for this purpose (Faraone 1985).

III. Mixed Models

Mixed models assume that a gene or genes of major effect act on a multifactorial background, with or without environmental effects. Very few of these studies have been done and the results are equivocal.

1.9.3 Anticipation and schizophrenia

The term anticipation is used to describe the increase in disease severity or decrease in age of onset in succeeding generations within families. Anticipation has been related to the expansion of trinucleotide repeats which cause at least ten inherited neurodegenerative/neuromuscular illnesses in humans including Huntington's disease, Fragile X syndrome and myotonic dystrophy. The phenomenon of anticipation has been reported in schizophrenia (Gorwood et al 1996) which has lead several groups to investigate the possibility of unstable trinucleotide repeats being aetiologically responsible for schizophrenia. The results from such studies are conflicting with Morris et al (1995), and O'Donovan (1995) reporting evidence for a modest excess of CAG repeat expansions in the genome of schizophrenia compared to controls as assessed by the Repeat Expansion Detection (RED) technique. The RED method utilises the ligase chain reaction. During a repetitive cycling reaction total genomic DNA serves as a template for the thermostable ligase to generate oligonucleotide multimers which represent expanded repeats. These can be separated on a polyacrylamide gel and visualised. When observed the CAG expansion was most marked in females, but others have reported no such expansion (Vincent et al 1996, Petronis et al 1996).

Family and twin data has been re-evaluated in the light of the discovery of unstable triplet repeats and Petronis and Kennedy (1995) concluded that such a mechanism competes well with the traditional multifactorial polygenic theory of schizophrenia with many deviations from Mendelian inheritance being explained by unstable DNA. Such a paradigm also may provide an explanation for several unclear issues in schizophrenia genetics such as discordance in monozygotic twins and the identical risk to discordant twin offspring of developing schizophrenia. Also the age of onset and perhaps even the phenotypic variation may be explained in terms of variation in the number of expanded repeats. In Huntington's disease the size of the expansion

of the CAG repeats appears to correlate to the age of onset which is decreased with increasing number of triplet repeats.

The hypothesis of triplet repeat expansions being involved in the aetiology of schizophrenia is particularly appealing for several reasons as they have been shown to be involved in the pathogenesis of several neurodegenerative diseases such as Huntington's disease in which there is selective neuropathology. The mechanism by which triplet repeat expansion causes such a selective neuropathology and whether the expansion is causative or a byproduct of the mechanism is unknown at this time but the elucidation of this mechanism will enhance our knowledge as to how such disease processes are selective and do not develop until later in life.

1.10 Environmental Aetiology:

Although there is strong evidence for a genetic component in schizophrenia, environmental factors, both biological and psychosocial, are likely to play an essential role in at least some cases of schizophrenia.

Several studies have implicated environmental factors in the aetiology of schizophrenia. Some of these environmental factors may act in concert with an underlying predisposition to schizophrenia and some may act alone to cause schizophrenia.

The small but significant seasonal birth effect, where schizophrenics are more likely to be born in the winter and spring, (see section 1.3.3) has lead many to postulate that some seasonally varying environmental factor which damages the developing foetus is responsible for at least some cases of schizophrenia. Infectious agents, environmental toxins and nutritional status are among the candidates proposed to explain this effect. Currently, infectious agents are most favoured since many display seasonal effects such as respiratory viral infections which are more prominent in the winter months. Viral infections, particularly the influenza virus, have been studied widely on the basis that infection at critical stages in gestation may have a detrimental neurodevelopmental effect which predisposes to schizophrenia in later life.

Mednick et al (1988) examined the effect of the epidemic of A2 influenza virus which affected Helsinki in 1957 and reported that those exposed to the epidemic during the second trimester of gestation were at elevated risk of developing schizophrenia. Studies by Kendell and Kemp (1989) failed to replicate this finding in

their Scottish sample but other studies such as that of O'Callaghan (1991) have replicated the finding in different populations.

Structural brain abnormalities such as ventricular enlargement, reduction in hippocampal volume, neuronal disorder and lack of gliosis have lead many to believe that schizophrenia is of neurodevelopmental origin. A viral insult in the second trimester of gestation which causes some developmental abnormality, is consistent with a neurodevelopmental hypothesis.

It is possible that an autoimmune response to infection could be responsible for schizophrenia (Knight et al 1992). It has been suggested that mothers who produce the optimal immune response to a viral infection such as the influenza virus may cause an autoimmune reaction in the foetus without causing any overt symptoms of influenza in the mother. Several studies have indirect evidence for infection in schizophrenia by demonstrating increased levels of immunoglobulins and cytomegalovirus antibody titre in the cerebrospinal fluid of schizophrenics (Wright et al 1993). In animals it has been demonstrated that antibodies directed against specific brain regions can be induced by viral infection such as group A Streptococcus and that anti influenza antibodies can cross react with human hippocampal, cortical and cerebellar tissues (Wright et al 1993). Increased levels of autoantibodies have also been demonstrated in the serum of schizophrenics and it is of interest that autoimmune disease affects females more often than males and that several studies have shown that the association between prenatal exposure to influenza epidemics and subsequent development of schizophrenia is statistically more significant in females. O'Callaghan (1994) did a study of 16 different viral infections strains and showed that none of these viruses were associated with increased incidence of schizophrenia indicating that the association seen with the influenza virus may be of significance in the aetiology of schizophrenia.

It has also been suggested that a new mutated virus could be responsible for schizophrenia but there is little evidence to support this.

Several studies show that complications occurring during birth or pregnancy are more common in schizophrenics that control subjects (Eagles et al 1991) particularly in male schizophrenics and are associated with earlier onset of schizophrenia symptoms (O'Callaghan et al 1992). Asphyxia, intracranial haemorrhage, ischaemia and malnutrition have been proposed as possible mechanisms which could cause damage leading to later development of

schizophrenia (Hemmings et al 1990, Murray et al 1992). It has been proposed that birth complications are a result of earlier events. Minor physical abnormalities are also found more commonly in several disorders including epilepsy and mental retardation but have also been reported in schizophrenia again particularly in males and are suggestive of an earlier developmental abnormality. It may be possible that obstetric complications could lead to the disruption of the blood brain barrier leading to increased susceptibility to viral or other pathogenic agents.

Schizophrenic individuals have also been found to have lower mean birth weights compared to their unaffected siblings and normal controls (Rifkin et al 1994). In monozygotic twins discordant for schizophrenia it been shown that the affected twin is significantly more likely to be lighter than their co-twin, although this finding has not been replicated (Rifkin et al 1994). Taken together these abnormalities suggest a neurodevelopmental influence in early to mid-gestation.

Other environmental influences on schizophrenia include psycho-social influences such as stress, socio-economic status and city residence. Schizophrenic individuals may be rendered more sensitive to environmental stress by the underlying abnormality which causes their schizophrenia. This may make them less able to cope with "stressful" events with which normal individuals have no difficulty. Stress is a general notion and what is stressful to one individual may not be to another. It is generally recognised that stress is a trigger of schizophrenia relapse and that major life events such as the death of a close relative, may precipitate episodes of schizophrenia rather than cause them (see earlier).

A higher rate of schizophrenia is found in inner cities compared to rural areas as shown by several studies including the Swedish city study of Lewis et al (1992). There may be several reasons for this which may incorporate some of the factors mentioned above, such as infectious agents being able to spread more easily in cities due to close living proximity and increased stress associated with urban living. Alternatively, the increase may be explained by "social drift" whereby affected individuals migrate to areas of a particular kind in which the social demands on them are less. Inner city areas provide less rigorous social demands on individuals whose coping skills are reduced and therefore preferential migration may occur. Pertinent to this argument is the high rate of schizophrenia seen among the lowest social class levels of urban populations (Kohn 1976). This is likely to be due to the downward social drift of affected individuals as a result of their disorder. Social

status may tell us more about how affected individuals fare in society as opposed to revealing any more information about the aetiology of the disorder.

It has been suggested that schizophrenia is more prevalent among African Caribbeans living in Britain than among other ethnic groups and the white population (Harrison et al 1988). This finding is controversial and in a similar study King et al (1994) reported that the prevalence of psychosis among all ethnic minority groups they studied was increased not just in Afro-Caribbeans. Although the initial report by Harrison lead some to suspect a biological basis for the difference in prevalence, King states that since the consistent finding of first or second generation migrants of any ethnic group may be at greater risk of a psychotic illness is suggestive of environmental factors. The greater stress which migrants suffer as a result of their migration, discrimination, cultural change and racism could play a role in the increased prevalence of psychotic disorders reported.

Another environmental contributor to the pathogenesis of schizophrenia which has been offered is that of abnormal metabolism of dietary products creating toxins which have a detrimental effect on brain function. This seems unlikely although in the genetically inherited homocystinuria (in which an enzyme defect preventing normal metabolism of the amino acid methionine occurs) psychiatric symptoms result from ingestion of dietary products containing the amino acid methionine (Hemmings 1990). Dohan 1982 suggests that in individuals who have genetic vulnerability to schizophrenia who eat grain products will develop schizophrenia. Dohan reported that the incidence of schizophrenia was highest in those countries which eat wheat and rye and lowest in those which consumed barley and rice. Abnormal metabolism of normal food stuff in situations of starvation may produce toxins which affect the brain (Susser 1995).

1.11 Hypothesis on the aetiology of schizophrenia

A plethora of hypothesis on the aetiology of schizophrenia have emerged incorporating genetic and/or environmental causes. Several hypothesis of the aetiology of schizophrenia incorporating major environmental components have been discussed in the previous section (1.10) such as viral, dietary and obstetric complications etc. and will not be discussed further in this section.

Neurotransmitter systems in schizophrenia

The role of the major neurotransmission pathways in the pathology of schizophrenia has been intensively investigated. The neurotransmitters, their receptors and other components involved in their production and functioning have all been scrutinised with the aim of defining a pathological role for them in schizophrenia. To date none have been conclusively found to be causally related to the pathogenesis of schizophrenia although they continue to be studied.

• Dopamine

The dopaminergic system regulates motor function and a number of cognitive and emotional functions which are transmitted through specific dopamine receptors.

The "dopamine hypothesis" is one of the oldest neurotransmitter theories and has been criticised extensively since its conception in 1973. The impetus for considering dopamine as an altered neurotransmitter involved in the pathogenesis of schizophrenia came from two observations. Firstly, that high doses of amphetamine, which acts by increasing dopamine release from nerve ending stores, can induce psychosis indistinguishable from that seen in acute paranoid schizophrenia. Secondly, that some antipsychotic drugs used to treat schizophrenia act via the blockade of dopamine receptors (D2 subclass) and the action of these drugs correlates particularly well with their clinical efficacy (although other neurotransmitter receptors may also be affected). As originally conceived the dopamine hypothesis posited that schizophrenic illness is a manifestation of a hyperdopaminergic state. There are however several difficulties with the hyperdopaminergia hypothesis. The many biochemical studies over the past two decades have failed to provide conclusive evidence of over activity of dopamine neurotransmission. The major dopamine metabolite, homovanillic acid (HVA) reflects, at least in part, dopamine function and has been measured in cerebral spinal fluid (CSF), plasma and urine. The results reported are conflicting, with some studies reporting elevated levels in schizophrenics compared to controls and others failing to demonstrate this. Direct measurements of HVA in post mortem brains have similarly failed to unequivocally show raised dopaminergic function in schizophrenia (Owen and Simpson 1994, Reynolds 1995). Post mortem brain studies did show that the number of dopamine D2 receptors in schizophrenics was elevated compared to normal control subjects. The source of this elevation is disputed as many of these patients received some form of drug treatment prior to

their death which may have caused up-regulation of these receptors. However, some studies done on drug free patients seem to confirm the increased number (Wong et al 1996).

Dopamine receptors belong to a super-family of receptors which exert their effect through guanine nucleotide binding protein or G proteins. Five subtypes of dopamine receptor are recognised; D1, D2, D3, D4 and D5. D1 receptors are coupled to adenylate cyclase and are located in the brain cortex. D2 receptors are negatively associated with adenylate cyclase and are found to be most prominent in the striatal and limbic structures. D3 is similar to D2 in terms of location and pharmacology although it generally has higher affinity for dopamine than D2. D4 is again D2-like and is found in the amygdala and frontal cortex whereas D5 resembles D1.

The antipsychotic drug clozapine is effective in treating some forms of schizophrenia resistant to treatment with classical neuroleptics and has fewer of the extrapyramidal side affects associated with these drugs. Clozapine binds to D2 receptors with low affinity, but seems to bind to D4 receptors with higher affinity (Meltzer 1994), indicating that D2 dopamine blockade is not the only pathway involved. All of the dopamine receptors have been investigated for polymorphisms to determine if susceptibility to schizophrenia could be linked to the presence of a particular functional variant. (This will be discussed further in section 1.14).

Hyperdopamine function correlates well with positive but not the negative symptoms of schizophrenia. Hypodopaminergic functioning in specific brain regions has been proposed. Davis et al (1991) hypothesised that schizophrenia was characterised by abnormally low prefrontal dopamine activity, causing the negative symptoms of schizophrenia, leading to excessive dopamine activity in mesolimbic dopamine neurones which causes the positive symptoms of schizophrenia. This hypothesis has received little acclaim.

Several functional studies have been done on the dopamine receptors using techniques such as positron emission tomography (PET) to visualise the receptors in specific regions of the brain in various groups of individuals. Recently Okubo et al (1997) showed by PET that binding of radioligand to the dopamine D1- like receptor (the D1 and D5 receptor subtypes) was reduced in the prefrontal cortex of schizophrenics and that this reduction was related to the severity of the negative symptoms seen in these individuals and to poor performance in the Wisconsin card sort test. Importantly this reduction was also seen in patients who had not been

exposed to antipsychotic drugs which can cause reduction in the levels of such receptors with chronic use. They proposed that dysfunction of the dopamine D1 receptor signalling in the prefrontal cortex may contribute to the negative symptoms and cognitive deficits seen in schizophrenia. However, as with many of these studies small numbers of individuals were used.

- **Serotonin**

Serotonin (5-hydroxytryptamine or 5-HT) was described as a neurotransmitter in 1950 and has been of continuing interest in relation to the pathogenesis of schizophrenia. The hallucinogenic drug LSD is an effective 5HT antagonist and this observation led to the hypothesis that schizophrenia might be related to a deficiency of 5HT. 5HT agonists have been shown to have an antipsychotic effect which is particularly effective at treating some of the negative symptoms of schizophrenia (Bleich et al 1988). Parallel studies to those of dopamine have been done on the major metabolite of 5HT - 5-hydroxyindolinic acid (5HIAA) in CSF, plasma and in post mortem brains. 5HIAA is reported to be decreased and correlated with ventricular enlargement in schizophrenia patients in some studies although, as with dopamine, the results are conflicting (Bleich et al 1988, Reynolds 1995).

5HT receptors have also been studied in brain tissue from schizophrenic patients. Several subtypes of 5HT receptor have been described (at least 14) and many have been implicated in the neurochemical pathology of schizophrenia, either directly or as a site of antipsychotic drug action (Roth 1994). Levels of the $5HT_2$ receptor in the brain have been studied, but results are inconsistent, with decreased densities being reported in several brain regions. Clozapine has a higher affinity for these sites than for dopamine D2 receptors and its administration seems to lead to rapid down regulation of these receptors and mixed D2 and $5HT_2$ receptor antagonists have been suggested for use in treating schizophrenia (Reynolds 1995). Clozapine and other antipsychotic agents have been found to have high affinity for several other 5HT receptor subtypes such as $5HT_6$ and $5HT_7$ which may mediate some of the clinical effects of these drugs. At present studies on these receptors are in their infancy. Further work is needed to clarify their role in the pathogenesis of schizophrenia.

- **Acetylcholine**

Anticholinergics may induce psychotic behaviour which may be related to the reciprocal nature of the dopamine and cholinergic systems in the brain. The

ubiquity of the cholinergic system in the limbic system of the brain and its involvement in a wide range of behavioural functions indicates that acetylcholine may play a role in psychosis. Tandon and Greden (1989) proposed that the negative symptoms of schizophrenia reflected hyperfunction of the cholinergic system in association with other neurotransmitter abnormalities, but there is little direct evidence in support of this hypothesis. The consideration given to the cholinergic system largely resulted from the effect of antimuscarinic drugs on relieving the acute extrapyramidal side effects of antipsychotic drug treatment. Clozapine has high affinity for muscarinic receptors and the reduced incidence of extrapyramidal side effects seen with use of this drug is considered to be related to this. Paucity of investigation of the cholinergic system in relation to pathogenesis of schizophrenia leaves may unanswered questions.

- **GABA** (γ-**A**mino**b**utyric **A**cid)

GABA is primarily synthesised by the inter-neurones in the cerebral cortex, hippocampus and limbic structures. These inter-neurones interact with several other neuronal systems including that of dopamine and glutamate. GABA is the major inhibitory neurotransmitter in the brain binding to $GABA_A$ receptor subtypes after synaptic release causing hyperpolarisation of the cellular membrane by stimulating chloride release and inhibiting neuronal activity (Coon et al 1994).

Roberts (1972) first proposed that GABA systems may be deficient in schizophrenia. GABA levels and GABA receptors have been reported to be reduced in some regions of the brain (hippocampus and temporal lobe structures), but as with the other neurotransmitters discussed thus far, there is contradictory evidence for this. Reduced GABA neurones and increased dopamine concentrations in the left amygdala have been reported (Reynolds 1995) and Reynolds speculates that this implies that neurones lost in schizophrenia could result in hyperactivity of dopaminergic neurones.

- **Glutamate**

Glutamate is ubiquitously distributed in the brain and is the primary transmitter of pyramidal neurones which are the principal excitatory neurones in the brain. Excitatory responses are mediated through a number of receptor subtypes which have a unique distribution being highest in the cortex, basal ganglia and hippocampus (Owen and Simpson 1994).

Initial interest in glutamate in terms of involvement in the pathology of schizophrenia was due to the finding that glutamate was reduced in the CSF of schizophrenic patients compared to controls, initiating the hypothesis that there was impaired function of the glutamatergic neurones in schizophrenia. This finding was not subsequently replicated but evidence that phencyclidine (PCP), which can induce psychosis in susceptible subjects, acts by blocking the ionotropic glutamate receptor n-methyl-D-aspartate (NMDA) ion channel renewed speculation of its involvement. PCP psychosis includes both the positive and negative symptoms of schizophrenia and is therefore thought to provide a better model than that of amphetamine induced psychosis as it more closely mimics the symptoms of schizophrenia. Further hypotheses about the role of glutamatergic dysfunction in schizophrenia have evolved from an understanding of the close interrelationship of dopamine and glutamate systems. Glutamate receptors including the NMDA receptors can stimulate dopamine release. It has been suggested by Toru et al (1994) that the relationship between dopamine and glutamate may explain the subclassification in so much as in those patients with primarily positive symptoms who are more likely to respond to antipsychotic treatment a hyperdopaminergic state exists. In those with negative symptoms refractory to antipsychotic treatment, abnormal neurotransmitter substances such as glutamate hypofunction may be involved.

Binding to the Kainate receptor subtype of glutamate receptor has been reported as decreased in the left hemisphere and increased in the frontal lobe consistent with loss of neurones on which the receptor is found and is in keeping with other reports of neurochemical indications of left temporal lobe dysfunction (Reynolds 1995).

Kerwin 1993 postulated that developmental abnormalities seen in the temporal lobe are due to abnormal assembly of microtubule associated proteins (MAP), the organisation and stability of which is influenced by glutamate. Reduction in glutamate or its receptors may cause such abnormal assembly.

- **Neuropeptides**

Several of the neuroactive peptides have been investigated in relation to pathophysiology of schizophrenia including;.

Neurotensin;

Neurotensin is a tridecapeptide which is found in several dopaminergic terminal regions including mesolimbic structures. Normal levels of neurotensin have been reported in general although there are some contradictions to this.

Cholecystokinin (CCK) and Somatostatin;

CCK is found in some mesolimbic dopamine neurones although it is primarily colocalized in cortical GABAergic neurones. Somatostatin interacts with many neurotransmitter systems but is similarly primarily co-localized with GABAergic neurones. Both CCK and somatostatin have been reported to be reduced in the hippocampus, amygdala and frontal cortex. It is unclear if this reflects loss of GABA neurones or specific neurotransmitter deficits.

- **Alpha-7 Nicotinic Receptor**

Schizophrenic individuals smoke at a higher prevalence rate (80%) than that seen in the general population (30%). Smoking appears to improve sensory gating, sustained attention and cognitive performance in patients with schizophrenia (Deleon 1996). Smoking may also decrease negative symptoms and extrapyramidal side effects and may also decrease the level of neuroleptic drugs in the blood thereby decreasing the doses required. These effects have been linked to the alpha-7 nicotinic receptor on chromosome 15. Using animal models Freedman et al (1994) showed that blockade of this receptor caused loss of the inhibitory gating response of auditory stimuli. Polymorphisms at the alpha-7 receptor have been identified and related to the sensory gating effect.

Cerebral Diabetes

Holden et al (1994) proposed a causal hypothesis for schizophreniform disorders of a form of cerebral diabetes. The underlying cerebral diabetic condition (so called since it is consistent with Positron Emission Tomography scans showing seriously abnormal glucose transport in schizophrenia) arises from decreased levels of P21 which is a GTP binding protein whose activity is regulated by GTPase activating proteins. This produces a reduction in levels of tetrahydrobioterin, a cofactor for the enzyme tyrosine hydroxylase and tryptophan hydroxylase which are involved in the biosynthetic pathway of dopamine and serotonin. This decrease compromises the activity of these two enzymes resulting in depression of dopamine and

serotoninergic neurotransmission and down regulation of glycometabolism. When such events occur at times of stress the blood brain barrier becomes more permeable to heterocyclic amines (present in cigarette smoke, pollution of various types and certain foodstuffs) the brain is "tipped over into a severe metabolic crisis which causes schizophrenia".

The membrane hypothesis

Prostaglandins (PG), especially prostaglandin E are known to modulate monoamine transmission. Activity of PGE may be decreased in schizophrenia resulting in increased dopamine and noradrenaline transmission (Neylan and Vankammen 1990). Abnormalities in essential fatty acids (EFA), which are precursors for PGs, form the basis of the membrane hypothesis (Horrobin and Huang 1993, Horrobin 1994). This proposes that the levels of EFA are decreased in the membranes of the red blood cells of schizophrenics (which are similar in composition to that of cells in the brain). The existence of defective acyl transferases, which insert certain types of EFAs into membranes and excessively active phospholipase A, which removes EFA from the membrane, are proposed to occur in the membrane hypothesis. Such abnormalities could result in disruption of the membranes of brain cells causing abnormalities in brain histology and in functioning of neurotransmitter systems (the components of which are all associated with membranes). Consistent with this hypothesis increased levels of phospholipase A have been reported in schizophrenics (Gattaz 1990) and neuroleptic drugs have been shown to inhibit phospholipase activity. Many of the epidemiological and pathological features of schizophrenia may be explained by such a hypothesis (Horrobin 1994).

Alcohol

Lohr and Bracha (1989) suggest that prenatal alcohol exposure could be a contributing factor to the development of schizophrenia in a small number of cases. Alcohol appears to be able to cause harm to the foetus when consumed even in small amounts. Enlarged ventricles have been reported in cases of foetal alcohol syndrome and hippocampal abnormalities are commonly seen in rodents exposed to alcohol. Lohr suggests that these data support the contention that prenatal alcohol exposure can cause brain structural abnormalities which may result in behavioural alterations. Late onset of symptoms may result from metabolic damage which only manifests after hormonal changes associated with puberty.

Epidemiological evidence that schizophrenia births peak in winter and that peak beer sales occur in late summer are taken to point to the critical time of prenatal exposure as being in early second trimester of gestation!

Handedness and the acquisition of language

Crow (1990) has hypothesised that schizophrenia results from a failure to develop the normal asymmetry of the brain represented by the presence of the language processing centre (Broca's region) in the temporal region on the left (in right handed people). Temporal lobe asymmetry is a relatively late evolutionary phenomenon and many believe that it is related to the acquisition of language skills.

Handedness is also believed to be derived from anatomical and physiological brain asymmetries and to be under the control of a "cerebral dominance" gene (Taylor and Amir 1995). The majority of the population are right handed (90%) but an excess of left handedness has been reported in schizophrenics (Green et al 1989). Post mortem studies and magnetic resonance imaging studies have also suggested that schizophrenic subjects lack pronounced asymmetry (reviewed by Maddox 1997). This evidence coupled with the fact that brain abnormalities are more often seen on the left side of the brain in schizophrenia has raised the possibility that handedness, cerebral asymmetry and psychotic disorder may depend in part on the same genetic locus. Crow has postulated that this hypothetical gene functions to induce the development of the cerebral asymmetry that usually puts Broca's area on the left and that a mutation would then account for its being on the right. Crow (1990) has also postulated that this gene lies in the pseudoautosomal region of the sex chromosomes (see section 1.14). Whilst unity hypotheses such as this are attractive, the evidence to hand (see earlier) suggests that schizophrenia is unlikely to be the result of a single gene disorder.

Neurodevelopmental abnormality

Structural brain imaging and epidemiology studies imply that schizophrenia may have its origins in foetal life.

Structural brain studies have indicated that abnormalities such as ventricular enlargement and decreased volume of the temporal lobe structures exist in some schizophrenic patients. Abnormal neuronal migration, greater pyramidal cell disarray (seen in some cases) and the lack of gliosis (glial cell reaction to neuronal inflammation or damage) suggest a developmental rather than a degenerative

process. Epidemiological indications that schizophrenia may be of developmental origin include the observation of premorbid clinical abnormalities in childhood (of personality and intellect), increased obstetric complications and increased winter births among schizophrenics the latter of which may indicate contact with an early environmental hazard such as a virus. All of these factors have been mentioned in more detail in previous sections.

There are many hypotheses as to the aetiology of schizophrenia but despite much experimentation, few have any convincing evidence to support them. Schizophrenia is undoubtedly a complex challenging disorder and is likely to result from the interaction of multiple genes of various effect sizes as well as environmental influences. Given this complexity and our relative ignorance about the genetic architecture of the higher central nervous system, finding susceptibility genes for schizophrenia is far from easy. New technologies and innovative methodologies have allowed progress to be made in positional cloning strategies in the field of psychiatric genetics as well as that of other complex disorders with a genetic basis.

1.12 Positional Cloning

The isolation of genes by positional cloning strategies is based purely on chromosomal location of the gene and requires no prior knowledge of its biochemical function. Positional cloning strategies have been successfully employed to elucidate genes involved in monogenic disorders such as cystic fibrosis (Rommens et al 1989) and Huntington's disease (The Huntington's disease collaborative research group 1993) but are now frequently being used to detect genes involved in the more common and challenging complex disorders including psychiatric illness.

Assignment of a disease gene to a chromosome and then sub-chromosomal region can be accomplished in several ways:

1.12.1 Molecular cytogenetics

Chromosomal rearrangements which cosegregate with diseases such as deletions, duplications, fragile sites and translocations provide an excellent starting point for positional cloning strategies and have proved extremely successful approaches particularly in cancer genetics where molecular events leading to the loss of a tumour supressor gene or the generation of gene fusion products can often be detected at the chromosomal level (Dallery et al 1995, Kerckaert et al 1993, Chen

1993). Duchenne muscular dystrophy was the first disorder to be successfully mapped using a chromosomal rearrangement (an X-autosome translocation, Greenstein et al 1977) and many others have since followed, including chronic Wilm's tumour and Fragile X syndrome (see Collins 1995 for a review).

Another type of rearrangement which has been associated with several diseases is triplet repeat expansions which can be readily detected by Southern blot analysis. Such expansions are found in Huntington's disease, Spinocerebellar and bulbar muscular atrophy (SMBA), Spinocerebellar ataxia type one (SCA1), Machado-Joseph disease and dentatorubral-pallidoluysian atrophy (DRPLA) and are of great interest as they are late onset neurodegenerative diseases. The precise method by which these triplet repeat expansions cause selective neuronal death and neurodegeneration is as yet unclear (Perutz 1996) (see section 1.9.3).

1.12.2 Linkage Analysis

For diseases where a large collection of affected families exists, the gene can be localised using genetic linkage analysis. Linkage occurs when a genetic marker (reliably measured characters that have simple mode of transmission and are polymorphic) and a disease gene lie in close proximity on the same chromosome such that they are found together more often in affected family members than would be expected by chance. Linked genes recombine less during meiosis than the expected rate of 50% stated by Mendel's law of independent assortment. The recombination fraction (θ) can be determined by dividing the number of individuals where there has been recombination by the total number of offspring and is proportional to the distance between the two loci (within certain limits). In practise statistical methods are used to determine the likelihood that the marker and disease gene are segregating independently. The LOD score (log of the odds) is the method usually employed to determine the likelihood of linkage at different recombination fractions. It is estimated for a number of recombination fractions and peaks where it reaches the maximum likelihood estimate of the recombination fraction. Linkage is assumed if the LOD score is 3 or greater (corresponding to an odds on linkage of 1000:1 or greater). The use of closely spaced polymorphic markers allows the region of linkage to be tightly defined as the markers which show the lowest frequency of recombination are those which flank the disease gene. The higher the number of meioses for study the better resolution is obtained. 100 meioses gives a resolution of about 1% recombination which equates to a genetic distance of

approximately 1 centimorgan (cM). In humans 1cM is roughly equivalent to 1Mb, however this relationship does not always hold true as recombination rates and physical distance vary throughout the genome.

Although linkage studies have proved successful in disorders where major gene effects occur such as Huntington's disease (Huntington's disease collaborative research group 1993) they have as yet been less successful in the more complex diseases such as schizophrenia. Linkage studies using maximum likelihood methods are based on precise specification of the genetic model (parametric analysis). In complex diseases genetic heterogeneity, multiple disease susceptibility loci, difficulties with precise phenotype definition, incomplete penetrance and non-Mendelian mode of inheritance all conspire to make linkage analysis difficult. Several methodological modifications and the use of larger sample sizes through international collaborations have helped to improve the power of linkage analysis in complex disorders. Non parametric linkage approaches, where the mode of transmission does not need to be specified, have been developed as an alternative to parametric linkage analysis. The affected sib pair (ASP) method is an approach which compares the number of alleles shared by descent between two or more affected sibs with the theoretical distribution of shared alleles under the assumption of no linkage (Penrose et al 1953). Evidence of linkage is inferred if affected pairs share more alleles at a given locus than would be statistically expected. This approach also has the advantage that it is easier to collect pairs of affected siblings than large multiplex families. The affected sib pair method is useful in detecting genes of moderate (as opposed to major) effect but the method suffers from low power and therefore very large sample sizes are required. Genetic heterogeneity, informativeness of the marker, the number of sibships and assortative mating all affect the power of this method and studies based on data from trios, with one or two affected members have been recommended. This approach has been successfully employed in complex diseases such as Alzheimer's disease (Baily-Wilson and Bamba 1993) and type 1 diabetes mellitus (Davies et al 1994).

Association studies are another non parametric test requiring no major assumptions other than the existence of a genetic contribution to the disorder. Association studies contrast to conventional linkage analysis in that they compare the frequency of marker alleles in a sample of unrelated affected subjects to that of a matched

normal population. The demonstration of a significantly different distribution of allelic variants in affected to unaffected individuals can result for several reasons;

I. Tight linkage disequilibrium between the marker and the disease susceptibility locus such that an association with a particular allele is maintained over many generations of recombination.

II. The marker locus may have some direct effect on susceptibility to disease as well as being detectable as a polymorphic marker (pleiotropy).

III. Population stratification - where the population contains a disease and marker allele more commonly than expected without a causal relationship between them. In a mixed population any trait present at a higher frequency in an ethnic group will show positive association with any allele that happens to be more common in that group (Lander and Schork 1994).

Population stratification can lead to false claims of association but may be reduced if the affected and control populations are ethnically well matched. Alternatively, the Haplotype Relative Risk (HRR) method of Falk and Rubinstein (1987) and variations of this method can be employed to overcome this difficulty. The control sample in this association study is derived from parental alleles not inherited by the proband, thus providing a perfect control for genetic background and avoiding the problem of population stratification.

Once a positive association is found it is often subjected to a further test known as the transmission disequilibrium test (TDT) to confirm linkage (Ewens and Spielman 1995). In this test a parent heterozygous for an associated allele A1 and a non associated allele A2 should transmit the A1 allele more often than A2 to affected offspring.

Association studies have several practical advantages over linkage studies, namely not requiring large numbers of affected family members which are difficult to obtain. They also don't require assumptions about the mode of inheritance and they have considerable power to detect genes of weak to moderate effect, unlike linkage where only genes of relatively major effect can be detected in practice. This is highlighted by the case of the insulin gene which showed strong association to the type 1 diabetes, but no linkage (this was also one of the first applications of the TDT test to sort out this problem)(Julier et al 1991). Linkage had been obscured because

of the substantial proportion of homozygous parents (non segregating and therefore not informative).

Association studies are currently conducted with markers close to or within a candidate region as they are limited to only being useful in testing the actual disease polymorphism or one in very tight linkage with it. Often this region may have been broadly implicated initially by linkage analysis.

It is now possible to do genome wide scans for linkage to determine regions where disease susceptibility genes may reside thanks to the great number of genetic polymorphic markers mapped throughout the genome, largely thanks to the efforts of the Human Genome Project. The results of such genomic scans and other linkage and association studies performed in schizophrenia will be discussed in section 1. 14.

1.12.3 QTL Linkage

Complex disorders such as schizophrenia are likely to result from multiple gene influences that, together with environmental variation, result in quantitative (continuous) distributions of phenotypes. Genes of various effect size in multiple gene systems that contribute to quantitative variation in a phenotype are often referred to as quantitative traits loci (QTL). Finding such QTLs has been possible in recent years by making use of animal models of various disorders and linkage analysis such as in the rat model of hypertension (Hilbert et al 1991). In this study linkage analysis (examining 49 DNA markers) in crosses between the stroke prone spontaneously hypertensive rat and the normotensive control strain (Wistar-Kyoto) led to the identification of two genes which significantly contributed to blood pressure variation in the F2 population (BP/SP-1: chromosome 10, and BP/SP-2; X chromosome) QTL linkage strategies have been utilised to study psychological traits such as "emotionality" of mice (Flint et al 1995). In this case the "open field activity" test was assessed to define emotionality of the mice. Subsequent inbreeding by brother sister mating over 30 generations resulted in each F2 mouse having a unique combination of alleles from the parent strains. The most and least active mice were examined for 84 DNA markers spread throughout the mouse chromosomes by linkage analysis. Three chromosomal regions were identified; chromosome 1, 12 and 15. A QTL on chromosome 15 was found to be primarily related to the open field activity.

Many differing QTL approaches are being employed to many domains of behaviour but statistical limitations mean that only reasonably large effects can be detected.

1.12.4 Synteny Homology

Comparative mapping, particularly of the mouse and human genome has proved useful in isolating disease genes involved in human disorders. The human and mouse genomes contain approximately 150 conserved regions with nearly identical gene content (Meisler 1996). Regions which are conserved between species make it possible to use gene identification and map position in one species to "predict" location in the other and to recognise true homologies between mouse mutants and human disease. The value of this technique is illustrated in the isolation of the gene responsible for Usher 1B syndrome, the most common cause of deaf-blindness in humans. The mouse deafness mutant *Shaker 1* and Usher 1B had previously been mapped to a conserved linkage group on mouse chromosome 7 and human chromosome 11q13 suggesting that they might be caused by a mutation in orthologous genes. The *shaker 1* gene was isolated in 1995 (Gibson et al) and was found to be an unconventional myosin, *myosin 7a*, expressed in the hair cells of the inner ear. This lead to the discovery of MYO7A gene in humans in which mutations were found in patients with Usher syndrome (Weil et al 1995).

It is also possible to use QTLs determined in mice as candidate QTLs in humans due to synteny homology. For example the QTL found in hypertensive rats BP/SP-1 could reside on human chromosome 17q which is homologous to the rat region of chromosome 10 to which this QTL was located. Similarly the QTL found on chromosome 1 in emotionality studies in mice is syntenic to the long arm of human chromosome 1. This allows these regions to be considered as candidate QTL regions for hypertension and fearfulness and anxiety respectively (Hilbert et al 1991, Flint et al 1995).

1.12.5 Mouse Models

Mouse models of human disease have a substantial role to play in the isolation of human genes. Induction of new mutants can be achieved using a variety of methods such as chemical mutagenesis but transgenic insertions and embryonic stem cell "knockouts" are providing increasing numbers of mouse mutations that are homologous to mutations in humans. The mouse is an ideal animal model due to its small size, short reproductive cycle and ease of genetic background manipulation

(i.e. ability to set up backcrosses, produce inbred strains etc) and is increasingly finding application in identification of disease genes and also determination of their function.

Such models are also finding application in determining the effects of genetic background and in isolating modifier genes such as *Mom-1.* The multiple intestinal neoplasia (Min) mouse is an excellent model for familial colon cancer. This mouse carries a mutant mouse *APC* gene and develops many intestinal adenomas. Genetic background has been shown to dramatically affect the tumour phenotype. This modification is due to a modifier gene, *Mom*-1 (modifier of Min), which maps to mouse chromosome 4 and can influence genetic variation in tumour number Luongo et al 1994, Moser et al 1992, Su et al 1992, MacPhee et al 1995, Dietrich et al 1993).

1.12.6 Physical mapping

Once the chromosomal region containing a disease gene has been narrowed down it may be possible to identify candidate genes which lie in that region based on known function or based on the availability of mapped Expressed Sequence Tag sites (EST's) (see later section). If such candidates do not exist then isolation of transcripts must be undertaken often the first step towards this is the production of a physical map of the region. The estimated size of the region will determine the number of markers required to produce a comprehensive physical map. Many markers have been produced and mapped as a consequence of the Human Genome Mapping Project and there is less need now to generate markers in the subintervals of those already available by techniques such as Alu-PCR from somatic cell hybrids and microdissection and microcloning of specific chromosomal regions.

The isolation and mapping of genetic markers is a critical step in positional cloning. The first class of polymorphic markers to be described were Restriction Fragment Length Polymorphism's (RFLP) which create or remove a restriction enzyme site such that digestion and hybridisation with an internal probe results in different band sizes in different chromosomes. These markers are not very informative as they tend only to have two alleles. Variable Number of Tandem Repeats (VNTR's)or minisatellites have also been used as markers overcoming the problem of insufficient allele number but are not evenly spread throughout the genome tending

to be clustered at telomeres. Microsatellites are more abundant (approximately every 30Kb), are more randomly distributed and are amenable to PCR amplification. Such markers have largely been superseded by sequence-Tagged Sites (STS's) which are short stretches (60-1000bp) of unique DNA sequence that can be detected by PCR. STS's have the advantage that only the sequence of the oligonucleotide primers for the PCR reaction is required to generate the STS and this information can be stored in electronic form therefore making the STS experimentally accessible to any laboratory without the need to obtain any cloned material. This therefore provides a common language on genetic maps allowing merging and integration of maps produced in different laboratories.

Human-rodent somatic cell hybrids have proved extremely useful in positional cloning allowing the assignment of new genetic markers and probes either to a particular chromosome or to specific regions on a chromosome, using Southern blot hybridisation or polymerase chain reaction (PCR).

Somatic cell hybrids are produced from the fusion of rodent and human cells and may be either monochromosomal (containing one human chromosome and rodent background), regional/deletion (containing fragments of human DNA in a rodent background), or radiation reduced (X-rays are used to break the human DNA into small fragments and the broken fragments are recovered in rodent cells). The technique of mapping using somatic cell hybrids is based on the loss or retention of specific chromosome or chromosome fragments in the hybrid. The hybrids are initially characterised using markers of known chromosomal position or by Fluorescent *in situ* hybridisation (FISH) thus allowing ordering of the hybrids to or on a chromosome. Monochromosomal panels first allow the marker/gene to be assigned to a specific chromosome then hybrid panels containing smaller pieces of that chromosome can be used to more precisely localise the marker/gene.

Somatic cell hybrids also provide some information about the order and distance of markers/genes. Segments of DNA which are further apart on a chromosome are more likely to be broken apart by radiation and segregate independently in the hybrid than if they were closer together. Additional information on gene families copy number and pseudogenes may also be gleaned by southern blot analysis of Somatic cell hybrids. Radiation hybrid mapping is currently being widely used in the Human Genome Mapping Project, providing a refinement of somatic cell hybrids which allow high resolution mapping.

Fluorescence *in situ* hybridisation (FISH) is widely used to determine map position and the relative order of DNA probes on chromosomes. The technique broadly consists of hybridising labelled probes to denatured target chromatin that has been fixed on microscope slides followed by visualisation using immunocytochemical staining procedures. Large genomic clones such as cosmids or Yeast Artificial Chromosomes (YACs) are required for this technique but it has the advantage of being able to detect gene amplifications, aneuploidy and subtle chromosomal rearrangements as well as being able to estimate distances between probes.

It is also useful to have a long range restriction map of the region being positionally cloned which can be achieved by Pulse Field Gel Electrophoresis (PFGE) (Brown and Bird 1986). Using this technique and a combination of rare cutter enzymes it is possible to refine physical distances between markers. As rare cutter enzymes occur in GC-rich sequences the location of CpG islands, which are likely landmarks for expressed genes, can be determined. The pulse field map of the region is also useful for determining the extent of coverage of overlapping clones in a contig in relation to actual genomic distance.

Once markers have been generated and mapped they can be used to generate a clone map of the region of interest. In long range physical mapping Yeast Artificial Chromosomes (YACs) are the cloning vector of choice. The large insert size of YACs (100-2000kb) means that fewer markers and clones are required to anchor and assemble the contig. YACs have been used to map several large chromosomal regions and have been successfully employed in many positional cloning strategies. There are several problems with YAC clones, the major one being the high percentage of clones which are chimaeric (clone insert is from two independent sources or chromosome regions and therefore does not represent a contiguous stretch of DNA). YAC clones are also quite unstable and may become rearranged or deleted over time.

Alternative cloning vectors to YAC's include bacterial based cloning systems such as bacteriophage P1cloning system (70-100kb insert size), P1 derived artificial chromosome (PAC - insert size 100-300kb) and the bacterial artificial chromosome (BAC - insert size up to 300kb) and cosmids (insert size 35-45kb). All these systems have the advantage of lower chimaerism, high transformation efficiency in library generation, ease of purification and are in general more stable than YACs. Such cloning systems are of use when dense ordered arrays of markers are present to anchor a contig.

Genomic clones which are isolated from a region are analysed for insert size and degree of overlap with other clones by marker content mapping and restriction enzyme mapping. The clones or there derivatives can be used for chromosome walking to close gaps in the contig and provide full coverage of the region. The genomic clones also provide DNA for the isolation of additional markers for use in techniques such as FISH, for generating sequence data and for gene identification.

1.12.7 Gene Identification

A major bottle neck in positional cloning has been the identification of transcribed sequences and several ingenious methods have evolved to address this issue;

- **Exon Trapping**

  Exon trapping is a method which facilitates the recovery of exons from random pieces of cloned DNA. The basis of the strategy is that during the retroviral life cycle, genomic sequences of non viral origin are correctly spliced and can be recovered as cDNA copy of the introduced segment. There are three distinct exon trapping methodologies that differ in the genomic target of interest. The original method described by Duyk et al (1990) was to capture isolated 3' splice sites residing within fragments of genomic DNA. Later approaches such as that of Buckler et al (1991) focused on either complete internal exons or entire 3' terminal exons (Krizman et al 1993). The choice of method is an important consideration as for example internal exon trapping results in usually a single exon of small size (20-200 bases) most of which contains useful protein coding information which can be used in database searches but is difficult to use for hybridisation screening of cDNA libraries to extend the sequence to full length. The 3' terminal exon trapping method generates sequence derived from the last exon of the gene and is generally much bigger in size but may consist mostly of 3'untranslated region which is less useful for database searching.

  Several exon trapping vectors are available. Most are plasmid based, containing a bacterial origin of replication, ampicillin resistance gene and a trapping cassette made up of a eukaryotic enhancer/promoter driving transcription of a two exon transcription unit with multiple cloning sites between the two exons. The first exon functions as a 5' terminal exon and the second as a 3' terminal exon capable of directing polyadenylation. Foreign fragments of genomic DNA are

inserted between the two exons and any internal exons present are trapped between the two vector exons by the cellular splicing mechanism.

Most DNA cloning vectors have been successfully employed as substrate in exon trapping such as plasmid, phage ,cosmids, BACs, YACs and even pooled clones. cDNA selection is currently the most favoured method for the isolation of gene sequences from a genomic region and has been successfully used in the case of the Huntington's disease gene (The Huntington's disease collaborative Research Group 1993). Exon trapping can not identify genes without introns or exons with cryptic splice sites and also has the disadvantage that it is only possible to use small DNA insert size.

- **cDNA selection methods**

Direct cDNA selection is an expression based gene identification technique that can rapidly identify cDNA within large genomic regions. The concept underlying direct selection is that large genomic regions such as YACs can be used to pull out cognate cDNAs from a mixture of cDNAs. This is sometimes called coincident cloning as it is the DNA which is coincident between the two mixtures which is extracted. If one of the DNA sources is genomic DNA from the region in which genes are required to be isolated and the other is cDNA from a particular tissue then the DNA which is coincident between the two resources will consist of gene fragments expressed in the specific tissue type.

Initially hybridisation of whole cDNA libraries on to immobilised genomic DNA (YAC or cosmid) was used after initial pre-blocking of high copy repeats (Lovett et al 1991, Parimoo et al 1991). This process was not particularly efficient despite the introduction of a PCR step to generate sufficient material for cloning and only gave 1000 fold enrichment (assessed by determining the relative frequency of a particular DNA sequence in the starting material as compared to the end product) which is insufficient to recover rare transcripts. A solution hybridisation reaction was then introduced in which both the genomic DNA and cDNA were in solution allowing the hybridisation conditions could be more easily controlled (Lovett 1994) and biotin-streptavidin capture systems were employed to separate the two resources. The basic procedure involves attaching a biotin moiety to the genomic resource during a PCR amplification step. The cDNA is pre-blocked for repeat sequences and both resources are denatured before being mixed together and heteroduplexes allowed to form between coincident

DNA. The denatured cDNA may be added to the genomic resource before or after it has been immobilised on streptavidin coated magnetic beads. Any coincident DNA is retained after stringent washing stages and can be eluted and either be used in further rounds to improve enrichment or cloned into a suitable vector after a further PCR amplification step and transformed into a bacterial host to construct a library which can then be screened.

The cDNAs used can be complex pools derived from several tissue types and the enrichment factors can be up to 10^3-10^6 (Brookes et al 1994) making the technique especially useful when searching for cDNAs that are expressed at very low levels in complex tissue sources. This technique is however prone to contamination with artefactual products. Pre-blocking of YAC genomic resource DNA with ribosomal yeast DNA and cDNA resource with human repetitive DNA helps to avoid contamination from these resources(e.g. Cot1 DNA).

End ligation coincident sequence cloning is a stringent variation of the coincident cloning technique which can enrich coincident sequences by 10^6 fold and is not confounded by the presence of repetitive elements (Brookes et al 1994). This technique is described in detail in the methods section 2.15. Basically, it allows the recovery of heteroduplexes that are sufficiently well matched internally to survive stringent washings and are perfectly base pair matched at their ends so as to allow the ligation of "capture oligonucleotides" in a brief high temperature ligation reaction which permits their recovery by amplification in a PCR reaction. The high temperature ligation step means that this method is more stringent than that described above and consequently less artefactual products should be obtained.

- **CpG islands;**

CpG islands are areas of the human genome which are G + C rich and, unlike the majority of vertebrate genomic DNA, don't show a reduced frequency of CpG dinucleotides and are not methylated (Bird 1986). CpG islands are usually located in or next to the 5' ends of genes. They commonly also contain the first exon of these genes. CpG islands can be detected using rare cutter enzymes or PCR techniques and used as hybridisation probes against cDNA libraries or Northern blots in order to identify the genes associated with them. However, not all genes are associated with CpG islands and for those that are associated

further analysis must be done in order to determine whether they represent transcribed sequences and if so, to isolate cDNAs representing these genes.

Screening of **northern blots** (mRNA derived from a tissue immobilised on a membrane) with genomic clones such as cosmids or phage (YACs are too complex to be used and must first be subcloned into either cosmids or phage) can be done to identify transcripts present in the tissue derived RNA. These transcripts must be present in sufficient amounts in order to be detected. The subclones which detect transcripts can then be used to screen cDNA libraries and are less complex probes than YACs. Similarly DNA subclones can be used to screen DNA from a variety of different animal species which have been immobilised on a membrane ("**zoo blotting**"). Detection of cross hybridisation between species may indicate the presence of sequences that have been evolutionarily conserved and therefore are likely to be essential for biological function. The main disadvantage of these techniques is their laborious nature when dealing with large genomic regions but they do have the advantage over techniques such as cDNA selection and exon trapping of not being as technically demanding.

- **Subtractive hybridisation techniques;**

Subtraction techniques have been used to isolate genes which are specific to a particular tissue source or developmental stage. The technique uses a target cDNA library, which is usually derived from the tissue where the gene is likely to be expressed, and a so called driver cDNA library who's function is to subtract out most of the ubiquitously expressed sequences. A typical experiment involves mixing two DNA samples that have been cut by restriction enzyme digestion, denaturing to form separate strands and reformation of heteroduplexes by reannealing of complementary strands. The tester DNA is usually mixed with a large excess of driver DNA to allow tester DNA fragments to predominantly form hybrids with the driver fragments. If the driver fragments are labelled with a hapten then separation of the driver and hybrid DNA can be achieved by affinity chromatography (Lisitsyn 1995). This process was found to be highly inefficient when trying to determine small differences between complex genomes and Representational Difference Analysis (RDA) was devised to overcome some of the initial limitations of the technique (Lisitsyn et al 1993).

RDA allows small differences between the sequences of two DNA populations to be determined. The efficiency is increased by initially cutting the driver and tester DNA with a restriction enzyme as previously but then ligating oligonucleotide adapters, allowing "whole genome" PCR preferentially of short fragments thereby reducing the complexity of the resources. Differences in the two DNAs results from restriction fragment length polymorphisms. The tester DNA is then ligated to oligonucleotide adapters before mixing with excess driver DNA after denaturation. On re-association target strand homodupexes with adapters ligated to the 5' ends, driver tester hybrids and driver homoduplexes. The target fragments can be selectively amplified after filling in of cohesive ends with Taq DNA polymerase before further rounds of enrichment after ligation of a new adapter to the tester sequence. RDA has been successfully applied to cancer genetics

- **Direct screening of cDNA libraries;**

YACs and other genomic clone inserts can be used to directly screen cDNA libraries removing the need to subclone and identifying multiple transcripts in a single screening. This process is however not straight forward requiring large amounts of DNA, long hybridisation times and giving high background as well as false positives due to the complex nature of the probe. Also poorly represented cDNA clones may be missed although normalisation of the library may circumvent this.

- **Direct sequencing, Software trapping and similarity searching;**

New automated sequencing methodologies are now making direct sequencing of relatively large genomic regions a realistic alternative to other gene finding techniques and will undoubtedly supersede such techniques in times to come. The vast array of database packages available to determine structural components of DNA sequences such as exon prediction programs (e.g. GRAIL, Xu et al 1996) are continually evolving becoming more efficient thanks to improved complex algorithms. Large scale sequencing strategies usually consist of shotgun cloning the region of interest into lambda phage, sequencing these clones usually by single pass automated sequencing, aligning the sequence into contigs using specially designed computer packages such as the Staden alignment package (Staden 1996) and closing gaps by a variety of techniques most usually by end cloning YACs. The sequencing effort currently being undertaken by the Human Genome

Project is using this strategy in its effort to sequence all the transcribed sequences in the human genome.

Once the sequence is obtained, rapid database searches such as (BLAST searches, Pearson and Lipman 1988, Aitschul et al 1990) are carried out to compare the sequence to all the known genes and expressed sequences already in the database. Such databases are currently growing at an enormous rate and hold large amounts of information not only on the sequence of the gene but of its expression pattern, mapping position and even known homologs.

♦ Expressed Sequence Tags (EST's)

Partial sequencing of cDNA clones to generate expressed sequence tags (EST's) is a rapid and powerful way to establish a profile of the expressed sequences in a particular tissue or cell type. Large scale cDNA approaches are currently being applied to a wide variety of tissues, organisms and developmental stages with a number of strategies being employed to yield maximum information from the libraries produced. The use of non-normalised libraries results in repeated sequencing of some cDNA clones but does yield information about the relative expression levels of active genes. Normalised libraries (which have been pre-screened to remove certain common sequences) enrich for clones which are more likely to be unique from one another (Sikela and Auffray 1993). The sequencing of clones also differs in that some labs sequence just the 5' ends of directionally cloned cDNA which is more likely to provide protein coding information useful for functional class assignment. Sequencing of 3' ends provides mostly 3' untranslated sequence which is less useful for functional assignments but since 3' UTR's are usually quite divergent they are more gene specific which often makes mapping easier.

The vast quantity of information produced from cDNA clone sequencing efforts is held in several databases including dbEST and GENBANK and is readily accessible through the world wide web (Boguski et al 1995).

Although there is a vast ever growing number of EST's held in these databases only a minority of them have been mapped to precise chromosomal locations, but this situation is likely to change as the Human Genome Project advances (see section 1.13).

1.12.8 Mutation detection

Once a gene has been identified in the most delimited region of DNA it becomes a candidate gene for the disease and its potential role in disease must be tested. This usually requires the demonstration that the disease is associated with a mutation in that gene which may be a single base pair change to more gross alterations such as deletions etc. While mutations can occur anywhere within the gene structure (introns and regulatory elements) most mutations to date have been found in the coding regions of genes. Gross gene rearrangements can be detected by pulse field gel electrophoresis but usually if the mutation is not the result of a major genic alteration more subtle mutations are demonstrated using techniques such as denaturing gradient gel electrophoresis, chemical cleavage of mismatches, single strand conformation polymorphism (SSCP) and direct DNA sequencing. SSCP has gained widespread use and is based on the fact that single strand DNA molecules adapt a secondary structure which dictates their migration rate through a gel in a non size dependent manner. A single base change alteration can substantially alter the conformation and therefore their migration through the gel is altered and can be seen as a band shift in comparison with normal migration. In this technique small fragments of 200-500bp (corresponding to exons usually) are amplified by PCR, denatured and separated in a non denaturing polyacrylamide gel. Demonstration of band shift will require sequencing of the DNA to establish the precise nature of the mutation but this technique is of great value when screening large genes for mutations (Orita et al 1989). New technologies such as the high throughput parallel array technique are currently being developed to allow refinement of mutation detection. These so called "DNA chips" consist of hundreds of thousands of oligonucleotide sequences contained on a ~2cm^3 chip (Southern 1996). The design of these chips varies depending on the requirements but if mutation (or sequence) analysis is required where every nucleotide must be interrogated then a complete series of four oligonucleotides spanning each position in the sequence will be designed, differing only in the identity of the central base. The relative intensity of hybridisation to each quartet of oligonucleotide probes reveals the identity of the central base. If a mutation has occurred then a different probe will light up because of the mismatch. Complex computer packages can convert this vast amount of information into comprehensible data and allow the rapid analysis of large numbers of samples at a relatively reasonable cost. DNA chips are being made commercially by companies such as Affymetrix and are now beginning to be used more widely in

the genomics field, having found applications in expression analysis and mapping as well as sequencing and mutation detection (Hacia et al 1996, Shoemaker et al 1996, DeRisi et al 1996).

Once a mutation has been detected it must be shown to be disease causing. Mutations which alter the structure of the protein product in some way which may alter its normal function is usually supportive evidence for its role in the disease. Ultimate proof that it is the disease causing mutation may come from transgenic studies by demonstrating that the mutation can cause the disease phenotype and that the normal form of the gene can correct the abnormal phenotype.
Other evidence may be accumulated from expression analysis from RT-PCR studies to see which specific tissues the gene is expressed in and also more specifically by RNA *in situ* which allows more detailed expression analysis in specific tissues. Evidence accrued from sequence information such as the presence of structural motifs which may give indications of function and homologies to other genes of known function either in humans or other organisms is often invaluable.
Tissue culture expression systems may also be employed to elucidate further information.

1.13 The Human Genome Project:

The Human Genome Mapping Project aims to produce a transcript map of the human genome by the year 2005 and ultimately to identify and establish a function for all of the estimated 50,000 to 100,000 genes in the human genome. The Human Genome Project also aims to provide new "infra structure" to genomic research by universal provision of data, material resources and technology that will improve the speed, cost and efficiency to aid biological research. The success of this process has been in the sharing of resources by scientific researchers across the world. At least 15 countries including Britain and America have programs for analysing the genomes of a variety of organisms including microbes and plants as well as the human genome. Many of the initial goals of these genome projects have already been realised with completed sequences of the entire genome of several species having been completed such as that of the bacteria *Escherichia coli* (1997) and yeast *Saccharomyces cerevisiae* (1993).
The initial stage of the human project was to produce relatively low resolution integrated cytogenetic, genetic and physical maps of all the human chromosomes.

The first low resolution genetic linkage map of the human genome was published in 1992 and since then many high resolution physical maps of various chromosome have been completed such as that for chromosome 7 and the X chromosome. Several region of the genome are currently being sequenced in large scale projects and an estimated 38.66Mb or 1.29% of the human genome has been sequenced to date (July 1997). The advent of new sequencing technology and strategies for large scale sequencing is rapidly accelerating the speed of this process and it is estimated that there is a world capability of sequencing 20Mb/year - a figure which will doubtless keep rising.

The Human Genome Project has had a remarkable effect on positional cloning strategies with thousands of markers being available throughout the genome (approx. every 100-200kb intervals) for linkage analysis. Once a region of linkage has been determined (depending on where it is in the genome) database searching can locate YACs from the region which can be ordered from centralised resource centres and cosmid clones may also be available. It is also possible that the region you are interested in will have had a number of EST's and other genes mapped to it which instantly provide candidate genes which can be further investigated. A large part of positional cloning can now be done by database searching and this fraction is set to increase as the Human Genome Project steams ahead.

As the scientific community hoped the Human Genome Project has had a phenomenal effect on the speed and ease with which human inherited diseases can be studied.

1.14 Previous attempts to find genes involved in schizophrenia;

Several strategies have been used in linkage and association studies to localise genes involved in schizophrenia. These include using *a priori* evidence for possible involvement of a region in schizophrenia such as a cytogenetic abnormality or the cosegregation of schizophrenia with another disease of known location. Alternatively linkage studies have been done by genome scanning using markers distributed throughout the whole genome.

Linkage Studies

- Chromosome 5

  In the wake of a report by Bassett et al (1988) of a partial trisomy chromosome 5 which cosegregated with schizophrenia in a single

Canadian-Chinese family, linkage to markers located between 5q11-q13 was reported by Sherrington et al (1988) in five Icelandic and two British pedigrees. A maximum LOD score of 6.49 was reported with broad spectrum diagnosis including alcoholism and schizophrenia spectrum disorders for a dominant gene model with high penetrance. However, a number of other studies using more informative markers in this region failed to reproduce this linkage (Kennedy et al 1988, St.Clair et al 1989, Crowe et al 1991, Macciardi et al 1992). The original linkage study was then extended using more informative highly polymorphic markers and the original LOD score was reduced substantially but remained positive at LOD 3 when the phenotype was restricted to schizophrenia and spectrum disorders. However, after new families were added to the original sample no linkage could be found and the original report was deemed to have been a false positive result.

- Chromosome 6

Wang et al (1995) and Straub et al (1995) both reported positive linkage results to chromosome 6. Wang et al performed linkage analysis on 186 Irish multiplex families and reported a multipoint LOD score of 3.9 at markers F13A1 and D6S260 (6P23) assuming locus heterogeneity and with broad disease definition. Straub analysed 265 Irish pedigrees and reported a LOD score off 3.51 assuming locus heterogeneity and intermediate phenotype definition with markers in the region of 6p24-22. Non-parametric analysis also yielded weaker but still positive findings suggesting a potential vulnerability locus for schizophrenia.

Several groups attempted to replicate these findings using the same markers as the original groups and varying results were obtained with some authors refuting linkage in their families at these markers (Mowry et al 1995, Gurling et al 1995) and others providing weakly positive results in support of the original findings (Antonarakis et al 1995). More recently Cao et al (1997) reported positive results for chromosome 6q using non-parametric allele sharing identical by descent and multipoint maximum likelihood score statistics but did not confirm the original 6p linkage finding.

- Chromosome 8

  Levinson et al (1996) reported a large collaborative study used informative pedigrees to study markers on chromosome 3, 6 and 8.

  Positive LOD scores were reported on chromosomes 6 and 8. The chromosome 8 maximum LOD score was 3.06 and the authors report this as being suggestive of a susceptibility locus for schizophrenia but that multicentre collaborative follow up studies are required for replication.

- Chromosome 11

  Several families have been reported where there is a balanced translocation on the long arm of chromosome 11q which appears to cosegregate with psychotic illness (Smith et al 1989 -t(9:11)(p22;q22.3), St. Clair et al 1990 - t(1;11)(q43;q21), Holland et al 1990 (6;11)).

  Several other lines of evidence make chromosome 11q interesting. Two independent pedigrees have been reported in which albinism cosegregates with schizophrenia (Baron 1976, Clark and Buckley 1989).This is of interest since the gene for oculo-cutaneous albinism is TYR (encoding tyrosinase) which maps to 11q14-21. Tyrosinase is of course also an important enzyme in the biosynthetic pathway of dopamine. Several candidate genes such as the dopamine D2 receptor gene also lie in this region of chromosome 11q making it a strong candidate region. However, results of linkage on this region of chromosome 11q have been largely negative, including those done on the dopamine D2 receptor and tyrosinase (Moises et al 1991,Nanko et al 1992, Wang et al 1993).

- Chromosome 22

  Pulver et al (1994) reported that chromosome 22q12-q13.1 showed positive linkage (LOD score 1.54) on a whole genome scan of patients from 39 systematically ascertained multiplex families. This study was followed up by a collaboration with several other groups using families from Ireland, USA, France and England to provide a combined replication sample of 256 families. No evidence of linkage could be found for the region 22q12-q13 leading the authors to conclude that the initial report was either false linkage or that the region confers susceptibility to

schizophrenia in only a small number of cases (Pulver et al 1996, Diehl et al 1994, Gill et al 1996, Laurent et al 1994). Further studies done on chromosome 22 have not yielded consistent results. Coon et al 1994 also reported positive linkage to chromosome 22 in a genomic scan done using 329 loci genotyped in 9 multiplex families but a similar genomic scan done by Barr et al (1994) in North Sweden did not reveal any such linkage. Polymeropoulos (1994) scanned 105 families with 10 chromosome 22 markers and found no evidence for linkage. Similarly Kalsi et al (1995) looked at two markers on chromosome 22 D22S274 and D22S283 (22q12-q13) in 23 UK and Icelandic families and no significant linkage was reported.

- Sex chromosomes

Collinge et al (1991) analysed a sample of 120 schizophrenic and schizoaffective sib pairs and reported that sibs shared marker alleles at DXYS14 locus more often than expected by random assortment. Positive preliminary results for DXYS20 and DXYS17 markers in the pseudoautosomal region were reported by d'Amato et al (1992) in 38 French pedigrees. The pseudoautosomal region is a region found on the distal ends of the short arms of both the X and Y chromosomes that undergoes crossing over during male meiosis. Alleles in this region are more easily transmitted to a same sex descendant and if dominant inheritance is assumed then affected offspring will tend to be of the same sex when the disease is inherited through the father (Leboyer 1992). Thus if the schizophrenia allele is on the fathers X chromosome then affected offspring are more likely to be female and if it is on the Y chromosome they will be male. Crow (1988) reported that same sex affected sibling were more likely to be concordant for schizophrenia. There are also reports that there is a high frequency of cytogenetic abnormalities (such as aneuploidies XXY, XXX) of sex chromosomes in schizophrenia (Crow 1988). This evidence was the basis of Crow's pseudoautosomal hypothesis for susceptibility to schizophrenia.

Asherson et al 1992 excluded linkage to the pseudoautosomal region (DXYS14) using six multiply affected families with schizophrenia. They did however find an excess of same sex siblings over mixed sex siblings

compared with that expected but conclude that this must arise through some alternative mechanism.

Outside of the pseudoautosomal region, Xq27-q28 was reported as showing linkage to bipolar disorder and was subsequently investigated for linkage to schizophrenia by DeLisi et al (1991 and 1994) who excluded linkage to a large part of this region with 10 multiplex families with schizophrenia.

1.15 Candidate genes

There have been many reports of studies done on loci thought to be candidate genes for schizophrenia.

- **Dopamine receptors;**

Despite many attempts to demonstrate linkage or association to any of the five dopamine receptors no consistent results have been reported (**D1**; Campion et al 1994. **D2**; Campion et al 1994, Nanko et al 1994. **D3;** Yang et al 1993, Nanko et al 1993, Nimgaonkar et al 1993, Nanko et al 1994. **D4**; Sommer et al 1993, Nothen et al 1994, Barr et al 1993 Shaikh et al 1994, Macciardi et al 1994, Seeman et al 1994, Lim et al 1994 (bipolar affective disorder)).

A slightly more consistent finding has been reported for the dopamine D3 receptor of excess homozygosity (Crocq et al 1992, Mant et al 1994). This finding has however not been replicated in other studies (Nothen et al 1993). Recently Nimgaonkar et al 1996 reported that allele 1 homozygosity at the D3RG locus increases the risk for and that this may also influence age of onset. They do also state that this is likely to account for a small amount of predisposition to schizophrenia and that liability may differ among Caucasian and African-Americans consistent with a polygenic/multifactorial mode of transmission. Durany et al (1996) have also recently reported homozygosity at the D3 receptor gene.

The dopamine transporter gene which functions to terminate dopamine function by the re-uptake into pre-synaptic terminals has also been widely studied but again no consistent results have been obtained though most do not favour linkage or an association to this region (Maier et al 1996, Li et al 1994).

Other Neurotransmitters;

- GABA

Mutations in the coding region of the β1 $GABA_A$ receptor have been sought by Coon et al (1994) using SSCP analysis. Direct sequencing of an SSCP variant revealed a C→G nucleotide transversion at codon 396 causing a histidine to glutamine substitution at a highly conserved site in the peptide but no association was found between this variant and schizophrenia.

- Serotonin (5HT)

Several authors have reported that the T102C polymorphism of the 5-hydroxytryptamine type 2a receptor gene (allele 2) is associated with schizophrenia (Williams et al 1996, Inayama et al 1996). Several subsequent studies failed to reproduce this finding (Chen et al 1997, Erdmann et al 1996, Malhotra et al 1996). Williams et al (1997) suggests this is due to the very small effect which requires a large sample size such as they used in their European multicentre association study of schizophrenia to improve the statistical power so as to detect the effect. They conducted a meta-analysis of 15 association studies which have been done on T102C 5HT2a and schizophrenia. Combining all the available data they conclude that there is evidence of an association between the T102C 5HT2a polymorphism and schizophrenia suggesting that the gene for the 5HT2a receptor or a locus in linkage disequilibrium with it confers susceptibility to schizophrenia (Williams 1997).

Porphobilinogen deaminase (PBGD);

Sanders et al (1991) reported an allelic association between schizophrenia and a polymorphism at the PBGD gene. PBGD gene is the site of pathogenic mutations which cause acute intermittent porphyria which can produce mental disturbances similar to those seen in schizophrenia and therefore is a possible candidate gene for schizophrenia. The association study done by Sanders et al revealed a relative risk for schizophrenia of more than 2 in individuals possessing the A2 allele. Homozygotes for the A2 allele may also have an earlier age of onset. This finding has not however been replicated (Owen et al 1992).

Dystrophin

Zatz et al (1993) reported a Brazilian family in which 4of 5 patients affected with Becker Muscular Dystrophy (BMD) also had a diagnosis of schizophrenia or related

spectrum disorders and suggested an association between the two conditions. This would imply that a locus for schizophrenia is located on the short arm of the X chromosome at Xp21 and the authors suggest that the psychiatric disorder may result from an abnormal expression of the dystrophin gene in the brain. However, there is little evidence to support this claim.

Beggs et al also reported that 2 out of a series of 35 patients with BMD had schizophrenia (Beggs et al 1991).

Huntingtin

Huntington's disease is a rare, severe inherited neurological disorder whose primary symptoms consist of involuntary choreiform movements and dementia. A small percentage of Huntingtons disease patients display symptoms of schizophrenia. Huntingtons disease results from an expansion of glutamine triplet repeats in the Huntingtons disease gene located on chromosome 4 and this has lead to the search for triplet repeat expansions in schizophrenia (see section 1.9.3). There has been one report of a schizophrenic individual who did not have Huntingtons disease having an expanded number of trinucleotide repeats within the Huntingtons disease gene (St Clair 1994) but Rubinsztein et al 1994 reported that no such expansion was found in the schizophrenic probands they studied.

Amyloid Precursor Protein (APP)

Mutations in the amyloid precursor protein (APP) have been reported as being involved in early onset Alzheimer's disease which is a late onset disease causing senile dementia. A large percentage of patients suffering from Alzheimer's disease display psychotic symptoms and this has lead to APP being considered as a candidate gene for schizophrenia. A point mutation in the APP gene has been reported in a schizophrenic individual (Jones et al 1992) but this was not confirmed in studies done on a further 100 schizophrenics by Jones et al or by Coon et al (1993) who looked at 86 schizophrenics. This variant is likely not to be associated with schizophrenia and is probably a rare polymorphism.

1.16 Project Background

In 1990 St Clair et al described a large Scottish pedigree (K26) in which schizophrenia and related disorders cosegregated with a balanced translocation t(1:11)(q43,q21).

This pedigree was detected by examining 282 pedigrees from the MRC Cytogenetics Registry with familial autosomal anomalies for mental illness using Research Diagnostic Criteria. 77 family members were included in the study (58 living and 19 dead). 23 members of the pedigree met Research Diagnostic Criteria (Spitzer et al 1987) for mental and/or behavioural disorders. The schizophrenic subjects also met DSM 111R criteria (American Psychiatric Association 1987).

Table 1.3: Number of individuals in pedigree K26 with mental illness

Disorder	Number of cases in K26 Pedigree
Schizophrenia	3
Schizoaffective Disorder	2
Major Depressive Disorder	6
Generalized anxiety Disorder	3
Minor Depressive Disorder	1
Alcoholism	3
Adolescent Psychiatric Disorder	5

No physical dysmorphisms or mental handicap were noted in this family. 34 of the 77 members of the family carried the t(1:11)(q43,q21) translocation and 16 of these 34 had a psychiatric diagnosis, the others were psychiatrically normal. Of the 43 family members who didn't have the translocation only 5 had a psychiatric diagnosis and none of these was of major mental illness (1 generalised anxiety, 1 minor depressive disorder and 3 alcoholism).

Linkage analysis was done and a range of gene frequencies and penetrances were implemented. A maximum LOD score (against chance linkage of the translocation with mental illness) of 2.19 was produced when the phenotype was restricted to schizophrenia and schizoaffective disorder and rose to 3.33 when recurrent major depression was included. A maximum LOD score of 4.34 was obtained when the phenotype included cases of adolescent conduct and emotional disorder. The LOD

score decreased to 1.37 however, when "uncertain phenotypes" such as generalised anxiety, minor depression and chronic alcoholism were included.

The identification of a cytogenetic abnormality which cosegregates with mental illness is an ideal starting point for defining the genetic component involved in the disease process in this family. With the hypothesis that a gene or genes involved in the psychiatric diagnosis in this family resides at or near the translocation breakpoint on chromosome 1 and/or 11, a positional cloning strategy has been employed to elucidate the relationship between the translocation and the psychiatric diagnosis in this family.

The existence of two previous reports of chromosome 11q translocations with chromosome 9 and 6 (Smith et al 1989 and Holland et al 1990) which had less clear cut cosegregation of mental illness, combined with the report of a large family in which schizophreniform psychosis cosegregated with tyrosinase-negative oculo-cutaneous albinism (Baron et al 1976) and several candidate genes, such as the dopamine D2 receptor gene, residing in close proximity to the translocation breakpoint on chromosome 11 indicated that chromosome 11q rather than chromosome 1 may be the region of interest.

Lymphoblastoid cell lines were set up from family members who carried the translocation and also from non carriers in the family. Genetic linkage, molecular analysis and *in situ* hybridisation studies identified Tyrosinase (TYR) and D11S388 anonymous marker as flanking the chromosome 11 breakpoint (Fletcher et al 1993). The distance between these two markers is estimated to be 7.8cM (NIH/CEPH Collaborative Mapping Group 1992).In order to generate more markers within this region to aid positional cloning the translocation-derived chromosomes were segregated into separate somatic cell hybrids using cell surface marker selection techniques (Fletcher et al 1993).Analysis of these hybrids confirmed that TYR was above the breakpoint and D11D388 was below the breakpoint. The chromosome 1 breakpoint was also defined with available markers and found to lie between markers D1S103 and D1S8.

Further mapping resources were obtained by X-irradiation of the chromosome 11 only hybrid J1C14 to produce four WJX irradiation hybrids (WJX 3, WJX 5, WJX 7 and WJX 11) which were assessed by PCR marker analysis and FISH with Human Cot1 DNA probes. To generate more markers from the region a microdissection and micro-cloning approach was adopted (Muir et al 1995). The breakpoint region on the der 1 chromosome is visible as an area of discontinuity under a phase

contrast microscope allowing its dissection. Two experiments were carried out both producing 160 fragments which were digested with *Eco*R1 and cloned into lambda phage. Over 1000 clones were derived from the two experiments and PCR of the inserts using primers directed at the vector sequences indicated that the cumulative insert size from both libraries (ignoring redundancy) was greater than 1Mb with the average insert size of the cloned being 1.5Kb. The libraries were screened for repeat sequences using $(GT)_{25}$ dinucleotide repeat probe to remove any repeat sequences. Of the 86 remaining clones 49 were assigned to chromosome 11 and 37 to chromosome 1. These clones were hybridised to Southern blots of the somatic cell hybrids previously made to define a series of intervals on chromosome 11(Fletcher et al 1993). FISH chromosome paints of pooled microdissection clones confirmed the fidelity of the microdissection clones to the breakpoint region.

5 of the microdissection clones which lay closest to the chromosome 11 breakpoint were used to screen ICI and ICRF YAC libraries (Anand et al 1990 and Lehrach et al,1990). 23 YACs were identified and constructed into a 3Mb contig using microdissection clone mapping and FISH for orientation. Confirmation of the contig and closure of gaps was achieved by end cloning of YACs. One of the 23 YACs was found to cross the translocation breakpoint (D0485) and this YAC was subsequently screened for brain expressed sequences using a cDNA selection technique, Coincident Sequence Cloning (CSC). This approach yielded an alpha tubulin pseudogene located 250Kb proximal to the breakpoint and six novel cDNA fragments which identify three novel genes located up to 700Kb proximal to the breakpoint. The genic nature of these cDNA fragments has been confirmed by sequence analysis, the presence of open reading frames and intron/exon boundaries and by expression studies including Northern blots and zoo blots (Devon et al 1997)

1.17 Aims of this thesis

This thesis has several aims:

1. To establish the precise location of two α actinin genes, ACTN 2 and ACTN 3, in relation to the translocation breakpoint by high resolution mapping techniques. These genes have been described as residing in the vicinity of the translocation breakpoints on chromosome 1 and 11 (Beggs et al 1992) and as such represent candidate genes for involvement in the psychiatric diagnosis seen in the translocation family whose candidacy must be further determined.

2. To establish the precise location of 20 ESTs with respect to the translocation breakpoint by high resolution mapping. These EST had previously been mapped to the general vicinity of the translocation breakpoint (Rosier et al 1995) and again represented candidate genes whose candidacy needed to be further determined initially by determining their proximity to the breakpoint.

3.To investigate linkage disequilibrium with a postulated schizophrenia susceptibility gene, close to the translocation breakpoint, in a case controlled allelic association study. This study will investigate two markers, in close proximity to the breakpoint on chromosome 1 and 11, in a random population of schizophrenics, unipolar depressives and normal controls.

4. To under take a search for genes in the region nearest the breakpoint on chromosome 1using a cDNA selection technique, coincident sequence cloning CSC.

5. To analyse clones from the CSC experiment and determine their genic nature and map location.

Chapter 2

Materials and Methods

2 Materials and Methods:

All chemical were supplied by BDH unless otherwise stated.
All solutions and methods were taken from or adapted with minor modifications from Sambrook et al (1989) unless otherwise stated.

2.1 Yeast Cell Culture and DNA Extraction

2.1.1 Media and solutions

All media were sterilised by autoclaving

AHC broth and agar

AHC is a rich selective medium which lacks uracil and tryptophan. It was used for selective growth of YAC recombinants prior to the isolation of YAC DNA.

In 1l water dissolve 1.7g yeast nitrogen base (without amino acids and NH_4SO_4) (Difco), 5g NH_4SO_4 and 10g casein hydrolysate (low salt). Adjust pH to 5.8. 17-20g of Bacto agar (Difco) were added to each litre of broth to make AHC agar. Autoclave. Add 50ml filter sterilised 40% glucose and 10ml of 2mg/ml adenine sulphate.

YPD broth and agar

YPD medium was used for growing yeast cultures.
In 1 litre water dissolve 20g bactopeptone (Difco), 5g yeast extract (Difco) and 5g NaCl. Add 1.5% agar for YPD agar. Autoclave. Add 40% sterile glucose.

SCE

In 500ml water dissolve 1M sorbitol, 0.1M Na_3Ci, 60mM EDTA, pH7.

2.1.2 Extraction of yeast DNA

A 5ml yeast cell culture was grown overnight in AHC at 30^0C and was centrifuged for 4 mins at 3000rpm (in a Sorvall T6000). The pellet was resuspended in 10ml 1M sorbitol, 0.1M EDTA solution, pH7.4 and transferred to a microfuge tube. 0.4ml of 2.5mg/ml zymolyase was added to break down the cell wall, and the tube incubated at 37°C for 2 hours. The tube was then centrifuged for 1 min at 11,000rpm, the

supernatant discarded and the pellet resuspended in 10ml of 50mM Tris.HCl pH7.4, 20mM EDTA.

0.5ml of 20% SDS was added, and mixed in carefully but thoroughly, before incubation at 65°C for 30 mins. The proteins were then precipitated by adding 4ml 5M KAc, pH5.4. The tube was then left on ice for 1hr, then centrifuged at 100,000rpm for 5 mins. The supernatant was transferred to a fresh tube and 1 volume of isopropanol was added. This was mixed and left at room temperature for 5 mins, before centrifuging at 11,000 rpm for at least 10 mins. The supernatant was then removed with a pastette, and the pellet left to air dry. The pellet was then resuspended in 6ml 10mM Tris 1mM EDTA, pH 7.4 (TE). RNA was removed by digesting with 30µl 10mg/ml pancreatic RNase (Sigma) at 37°C for 30 mins. The DNA was then precipitated by adding 300µl 5M NaCl and 18ml 100% ethanol, and then pelleted by centrifugation. The supernatant was then removed, the pellet was washed in 500µl of cold 70% ethanol and air dried. The pellet was then resuspended in 1ml TE, pH7.4.

2.2 Bacterial Cell Culture and Plasmid DNA Preparation

2.2.1 Media and solutions

All media were sterilised by autoclaving.

Terrific broth

In 900ml water dissolve 12g tryptone (Difco), 24g yeast extract (Difco), 4g glycerol. Autoclave then add 100ml autoclaved phosphate solution (0.1M KH_2PO_4 & 0.72M K_2HPO_4).

L-Broth and agar

In 1L of water dissolve 2.46g MgSO4, 10g tryptone (Difco), 5g yeast extract (Difco) and 10g NaCl. 15g agar (oxoid Ltd) was added per Litre broth for L-agar.

Ampicillin (Sigma)

Ampicillin was added to agar and broth in order to select for bacteria carrying plasmids which confer resistance to this antibiotic.

Make stock at 50mg/ml in dH_20, filter sterilise. Store at -20°C.

Use at final concentration 50μg/ml.

5-Bromo-4-Chloro-3-Indolyl β-D- Galactopyranoside (X-Gal) (Sigma)

X-Gal acts as a substrate for β-Galactosidase. It was added to agar for the growth of plasmids carrying blue/white colour selection (see section 2.5.6).

Make stock at 20mg/ml in DMF (Sigma). Store protected from light at -20°C.

Use at final concentration of 40μg/ml.

Isopropyl β-D- Thiogalactopyranoside (IPTG) (Sigma)

IPTG is a derepressor of the Lac operon. It was added to agar for the growth of plasmids carrying blue/white colour selection.

Make stock at 100mM. Store protected from light at -20°C.

Use at final concentration of 0.5mM.

Hogness Solution

3.6mM K_2HPO_4, 1.3mM KH_2PO_4, 2mM $Na_3Ci.2H_20$, 1mM $MgSO_4.7H_20$, 4.4% glycerol.

Make stock at 10x concentration. Filter sterilise.

GTE

50mM glucose, 10mM EDTA, 25mM Tris.HCl (pH 8)

Use autoclaved Tris and EDTA. Sterile filter the glucose.

2.2.2 Growing bacterial cells on agar plates

The desired volume (up to ~200μl) of bacterial cells is pipetted onto the surface of the L-agar, and is then spread by a sterile bent glass rod until the liquid has thoroughly soaked into the agar. (If more cells than are contained in 200μl are required, the cells may be concentrated by 1min centrifugation in a microfuge followed by resuspension of the pellet in a smaller volume.) The plates are then inverted and incubated for 12-16 hours at 37°C. Cells were viable from plates stored at 4°C for several weeks.

2.2.3 Frozen stocks of bacterial colonies

The selected colony was re-streaked on a fresh L-agar plus ampicillin plate, then a single colony from the new plate was used to inoculate 5ml of Terrific broth plus ampicillin. After incubation at 37°C for approximately 16hrs, a 1/10 volume of 10x Hogness solution was added to 1ml of the culture in a standard 1.5ml eppendorf tube (Treff). This was then frozen and stored at -70°C.

To reactivate cells from frozen stocks, a plastic loop was passed across the surface of the frozen culture and sufficient cells would be picked up to spread on an L-agar plate and grow at 37°C.

2.2.4 Use of colony picker

A colony picking robot (Hybaid) was used according to the manufacturer's instructions to spot the CSC product library colonies onto filters and in the preparation of their frozen stocks.

2.2.5 Extraction of plasmid DNA (small scale)

The method used was a modification of the alkaline lysis method of Jones and Schofield (1990).

The desired colony on an agar plate was used to inoculate 6ml of Terrific broth plus ampicillin and was grown at 37°C overnight with continuous shaking at 250rpm. The culture was then centrifuged for 1min in a standard 1.5ml eppendorf tube, adding 1.5ml of the culture at a time and retaining the pellet. The pellet was then resuspended in 200µl GTE (50mM glucose, 10mM EDTA, 25mM Tris.HCl (pH 8) and left on ice for 5mins. Next, to lyse the bacterial cell wall and denature the DNA, 500µl of freshly prepared 0.2M NaOH/1% SDS was added, mixed thoroughly by vortexing and left on ice for 5mins. To neutralise, 250µl of 3M KAc (pH 4.8) was then added, mixed thoroughly as before and left on ice for 10 mins.

To remove the cellular debris, the tube was centrifuged twice for 10 mins at RT at 11,000rpm and in each case the supernatant was retained. RNase A (Sigma) was then added at a final concentration of 20µg/ml and left at 37°C for 10mins. The

supernatant was extracted twice with 200µl chloroform (Fisons), mixing the layers thoroughly by vortexing and then separating them by centrifugation for 1min.

The DNA was then precipitated by adding an equivalent volume of 100% isopropanol (Fisons) and immediately centrifuging for 10 mins at room temperature. The supernatant was discarded and the pellet then washed briefly in 500µl 70% ethanol and dried under vacuum.

2.2.6 Precipitation with polyethylene glycol (PEG)

The DNA prepared for ABI sequencing was subjected to a further precipitation. The pellet was resuspended in 33.6µl dH_20, then 6.4µl NaCl and 40µl autoclaved 13% PEG_{8000} were added. After thorough mixing, the tube was incubated for 20 mins on ice and then the DNA was precipitated by centrifugation for 15 mins at 4°C. The supernatant was then removed and the pellet dried under vacuum. The pellet was then resuspended in 20µl of dH_20. To check the concentration of plasmid DNA recovered, 1µl of the 20µl was run on a 0.8% agarose gel, and the intensity of the bands compared with standards.

2.2.7 Extraction of plasmid DNA (large scale)

The selected bacterial colony was grown overnight at 37°C, with shaking at 250rpm, in 4ml Terrific broth plus ampicillin. This 4ml was then used to inoculate a further 400ml of Terrific broth for an overnight culture. This culture was then centrifuged in 50ml volumes at 4000rpm for 20mins at 4°C. The pellet was then resuspended in 16ml GTE and left for 5mins on ice. Next, 32ml of freshly prepared 0.2M NaOH/1% SDS was added, mixed thoroughly by vortexing and left for 5mins on ice, before adding 16ml 3M KAc (pH 4.8), again mixing and leaving for 10 mins on ice.

The tubes were then centrifuged at 4000rpm for 20 mins at 4°C and the supernatant retained. This supernatant was filtered under gravity through 4 layers of gauze into fresh centrifuge bottles. 36ml of isopropanol was then added and the bottles spun immediately at 4000rpm for 10 mins at 4°C. The supernatant was discarded and the pellet was washed with 20ml of 70% ethanol and centrifuged at 4000rpm for 10 mins at 4°C. The supernatant was once again discarded and the bottles drained. The pellet was resuspended in 2.5ml TE (pH 8) and transferred to a Universal tube.

3.55g CsCl and 200µl of EtBr (5mg/ml stock) (Boehringer Mannheim) were added and the solution transferred to a 'quickseal' tube (Beckmann). A balance tube was made up if necessary with TE, CsCl and EtBr. The tube was then spun for 20hrs at 80,000rpm at 20°C in a fixed angle rotor. The plasmid band was then carefully extracted through the side of the tube with a needle and syringe and placed in a 15ml tube. The plasmid DNA was then extracted three times with a 0.25x volume of CsCl-saturated isopropanol. Water was then added to increase the volume approximately twice, and then the DNA was ethanol precipitated. The resulting pellet was resuspended in 1ml water and the concentration of plasmid DNA checked by running 1µl on a 1% agarose gel and comparing the band intensities with standards.

2.3 Extraction of cosmid and PAC DNA

Cosmid and PAC DNA was extracted as described above (2.2.7) large scale plasmid preparation.

2.4 Extraction of genomic DNA from Cultured Cells

Frozen cell pellets were allowed to thaw before being resuspended in TNE (10mM Tris, pH 8.0; 1mM ethylenediaminetetra-acetic acid (EDTA); 150mM NaCl). The cells were then lysed by the addition of 2.5ml lysis buffer (0.5% sodium dodececyl sulphate (SDS);150mM NaCl; 100mM Tris.HCl, pH 8; 100mM EDTA). RNA was digested by the addition of 300µg RNase A (Sigma) and incubated at 37^0C for 15 minutes. Proteins were then digested by the addition of Proteinase K (Boehringer Mannheim) at 0.5mg/ml with incubation overnight at 50^0C. An equal volume of water saturated phenol was then added to the extracts and mixed thoroughly before being centrifuged at 3000rpm for 5 minutes. The aqueous layer was removed and extracted with an equal volume of 1:1 Phenol/chloroform in the same way as described above. The aqueous layer was removed and then extracted with an equal volume of chloroform. The DNA was then precipitated from the aqueous layer with 0.5 volumes of 7.5M ammonium acetate and two volumes of ethanol. The DNA was spooled out of the solution onto a glass rod and allowed to air dry. The DNA was then washed in 70% ethanol and again allowed to air dry. The DNA was then dissolved in 200µl TE (10mM Tris.HCl pH 7.5; 1mM EDTA).

2.5 Cloning of DNA Molecules into Plasmid Vectors

2.5.1 Plasmid vector

The plasmid vector used was pBluescribe (pBS; Stratagene) which is a 2746bp plasmid derived from pUC19 (Yanisch-Perron et al, 1985). It contains an ampicillin resistance gene and a multiple cloning site flanked by T3 and T7 promoters.

All the DNA molecules cloned possessed cohesive termini, which were compatible with a restriction enzyme site in the multiple cloning site of pBS. The digested insert molecules were therefore directly ligated into pBS linearised with the appropriate enzyme. So as to prevent recircularisation of the plasmid during ligation the 5'-phosphate groups from the vector were firstly removed by phosphatase treatment. The circular plasmid containing an insert of DNA was then used to transform competent cells.

2.5.2 Strain of bacteria used

The competent cells were made from *E.coli*, strain XL1-Blue (Stratagene). The genotype of XL1-Blue cells is: recA1 endA1 gyrA96 thi-1 hsdR17 supE4 relA1 lac [F' proAB $lacI^q$ ZΔM15 Tn10 (Tet^r)]

2.5.3 Preparation of competent cells for electro-transformation

XL1-Blue bacterial cells were taken from a frozen stock using a sterile loop and streaked out onto a fresh L-agar plate and grown at 37°C overnight. One of the resulting colonies was then used to inoculate 10mls of Terrific broth plus tetracycline, which was left at 37°C in a shaking incubator overnight. This whole culture was then used to inoculate 1l of Terrific broth, and this was grown shaking at 37°C. The absorbance at 600nm of the culture was continually checked on a spectrophotometer.

The cells were harvested when the Abs_{600} was between 0.5 and 1.0, which is still in the log phase of growth. The flask containing the culture was then chilled on ice for 15-30 mins then the contents were centrifuged 50ml Falcon tubes (which had been left on ice) for 15 mins at 4000rpm at 4°C. The supernatant was discarded and the pellets resuspended in a total of 1l of sterile water before centrifugation as before. The pellets were resuspended in a total of 500ml sterile water and re-centrifuged. Pellets were then resuspended in 20ml of 10% glycerol and centrifuged again. The

final resuspension was in 2-3ml of 10% glycerol. This cell suspension was then aliquoted into 40µl volumes in 1.5ml eppendorf tubes and snap frozen on dry ice plus methanol (Fisons). The aliquots were stored at -70°C.

2.5.4 Test transformation of competent cells

In order to quantify the competence of the cells prepared as above, they were subjected to a test transformation using 1ng of pBS vector alone as the transforming DNA. Cells transformed with pBS contain an ampicillin resistance gene and can therefore be selected for on L-agar plus ampicillin plates. The competence of the cells is measured in colony forming units (cfu's) per µg of transformed DNA. Values of 10^7 to $5x10^8$ cfu's/µg pBS were routinely achieved.

2.5.5 Electro-transformation of competent cells

An aliquot of competent cells was thawed on ice, and the transforming DNA (usually 1µl of the ligation reaction) was added, mixed with the cells and left on ice for 1min. This mixture was then transferred to an ice-cold cuvette (Flowgen) and subjected to a pulse of 2.5kV in a BioRad Gene Pulser. The cells were swiftly mixed with 1ml of Terrific broth, transferred to an eppendorf tube and left at 37°C for 1hr to enable the cells to begin to express the ampicillin resistance gene conferred by the transformed plasmid. Aliquots of several different volumes were then spread onto L-agar plus ampicillin plates and incubated overnight at 37°C.

2.5.6 Selection for colonies that contain recombinant plasmids

Selection of colonies containing recombinant plasmids was achieved using blue/white colour selection. The polylinker of the vector pBS interrupts the β-galactosidase producing gene, lac Z. This disruption is in-frame and results in a harmless insertion of a few amino acids into the β-galactosidase gene. Expression of this gene within bacterial cells can result in the production of a blue colour if the medium on which the cells are grown contains the chromogenic substance X-Gal and the derepressor of the Lac operon, IPTG. If however the polylinker site in the plasmid is interrupted by an insert of foreign DNA, a functional β-galactosidase enzyme cannot be transcribed and hence the resultant colonies are white.

2.6 Electrophoretic analysis of DNA

2.6.1 Electrophoresis solutions

Stock solutions of electrophoresis buffers were made:

20x TAE

0.8M Tris.HCl, pH 8.0; 20mM EDTA; 0.4M acetic acid.

20x TBE

1M Tris.HCl, pH8.0; 20mM EDTA; 1M boric acid, pH8.3.

Electrophoresis gels were run in 1X electrophoresis buffer.

10x DNA Loading Buffer

20% Ficoll (Pharmacia), 100mM EDTA, orange G (Sigma).

6% Polyacrylamide Gel (for ALF automated genotyping)

Gels were made up with 6% Hydrolink (Long Ranger™, AT Biochemicals Ltd) in 0.6X TBE (0.06 M Tris base, 0.05M boric acid, 0.6mM EDTA)containing 7 M urea.

2.6.2 Size markers used in Gel Electrophoresis

250-500ng of the appropriate size marker was used per gel. Size markers used were:

a) λ DNA digested with HindIII (Boehringer Mannheim)

b) ϕX174 digested with HaeIII (Boehringer Mannheim)

c) 1kb ladder (Gibco BRL)

The sizes of the bands in λHind III, ϕX Hae III and the 1kb ladder are listed in Table 2.1.

Table 2.1: Band sizes of markers λ Hind III, ϕX Hae III and 1kb ladder.

λ Hind III (kb)	ϕX Hae III (bp)	1kb ladder (bp)
23	1353	12216
9.5	1078	11198
6.5	872	10180

4	603	9162
2.3	310	8144
2.0	281	7126
0.5	271	6108
	234	5090
	194	4072
	118	3054
	72	2036
		1636
		1018
		517
		506
		396
		344
		295
		220
		201
		154
		134
		75

2.6.3 Agarose gel electrophoresis

DNA molecules were separated according to their size on horizontal agarose medium EEO (Sigma) gels. The percentage of agarose used to make the gel depended on the size range of the DNA molecules to be resolved. Digested genomic DNA or plasmid DNA was commonly run on 0.8-1% agarose gels, whereas smaller fragments, such as most PCR products, were run on 1-2% agarose gels. All agarose gels were made with and run in either 1XTAE or 1XTBE. To stain the DNA, ethidium bromide was added to all agarose gels and to the running buffer at a concentration of 250μg/ml of buffer. 1X loading buffer was added to the DNA prior to loading the sample on the gel. Gels were run at 25-80V depending on resolution and run-time required.

DNA fragments were visualised on a mid range UV transilluminator and photographed using a video copy processor (Mitsubishi).

2.6.4 Polyacrylamide Gel Electrophoresis on Automated Laser Fluorescent (ALF) Sequencing gels for allele genotyping

Electrophoresis was carried out on a 40-lane automated laser fluorescence (ALF) sequencer (Pharmacia) using 6% polyacrylamide gels in 0.6X TBE (0.06 M Tris base, 0.05M boric acid, 0.6mM EDTA) and run in the same buffer at constant power (55W) and constant temperature (50°C).

After the PCR reaction the products from both reactions (D11S931 and D1S1621), for the same patient, were run in the same lane on the ALF sequencer (since the products are of different size) with appropriate size markers and gold standards. Automated genotyping was carried out using 'ALF' manager and 'ALP' software.(He *et al* 1995, Mansfield *et al* 1994). All gel results were also checked manually.

2.7 Purification and Concentration of DNA

2.7.1 Ethanol precipitation

To concentrate DNA and remove salts, a 1/10 volume of 2M NaAc, pH 5.5 was added to the DNA solution, followed by 2-3 volumes of 100% ethanol at -20°C. The contents of the tube were mixed and then chilled at -70°C for 1hour to overnight. The tube was then centrifuged at 11,000rpm for 15 mins at 4°C. The supernatant was poured off, and the pellet dried under vacuum. The pellet was then resuspended in the desired volume of dH_2O.

2.7.2 Phenol/chloroform based extraction

To remove proteins from DNA samples, the DNA was extracted with organic solutions. Two methods were used:

1)Phenol/ chloroform extraction:

DNA was extracted twice with a phenol/water/chloroform mixture (Applied Biosystems). A 1/10 volume of this organic mixture was added to the DNA solution to be extracted, vortexed vigorously and then centrifugation at 11,000rpm for 2 mins. The upper, aqueous layer was then removed carefully, avoiding the

precipitated protein at the interface between the two layers. The DNA was then recovered by ethanol precipitation.

2) Chloroform/isoamylalcohol/ether:

DNA was extracted first with a 1/10 volume mixture of chloroform (Fisons)/isoamylalcohol (Sigma) in a 24:1 ratio. A second extraction was then performed with an equal volume of water-saturated ether (Fisons). After separation of the aqueous layer from the ether, the tube was left open to the air for 10 mins to allow evaporation of any residual ether, and then DNA was ethanol precipitated.

2.7.3 Drop dialysis

DNA solutions were drop dialysed to remove salts. A 0.025μm drop dialysis filter (Millipore) was placed on the surface of a pool of sterile water in a Petri dish. After 5 mins, the DNA solution (maximum 20μl) was placed as a drop onto the surface of the filter and left for 30 mins, during which the concentration of salts in the sample equilibrates with the water. The drop was then 'sucked back up' into a clean tube.

2.7.4 Purification of DNA from agarose gels

DNA fragments were run in low melting point agarose (Ultrapure LMP agarose, Gibco BRL) in 1x TAE. Gels were viewed on a mid range UV transilluminator and the required fragment was excised using a sterile scalpel blade as quickly as possible to minimise nicking of the DNA. Care was taken to ensure that a minimum of agarose was excised with the required DNA band.

Isolation of DNA from gel slices was by agarase treatment. The gel slice was diluted with water to decrease the percentage agarose, then 1/25 volume of 25x agarase buffer (Boehringer Mannheim) was added. Large gel slices were split into aliquots of ~100μl each. The gel slices were then melted at 70°C for 10 mins, during which they were vortexed and spun down several times to ensure they were properly melted. The slices was left to cool to 37°C, then 1μl (0.5-1unit) agarase (Boehringer Mannheim) per 100μl was added. The reaction was left to proceed for 2hrs at 37°C, then the DNA was ethanol precipitated at -70°C from 1hr to overnight.

2.7.5 Geneclean (Bio 101 Ltd)

The Geneclean Kit was used to purify DNA from agarose gel slices in accordance with manufacturers instructions.

2.8 Transfer of DNA to membranes

2.8.1 Southern Blot transfer

DNA was transferred from gels to nylon membranes by capillary blotting. This method was adapted from Southern (1975).

Gels were photographed next to a ruler to allow for future sizing of DNA fragments.

Firstly the DNA was denatured by gently shaking the gel in denaturing solution (0.5M NaOH, 1.5M NaCl) for 30-45 minutes. The gel was then neutralised by gently shaking in neutralising solution (1M Tris.HCl, 2M NaCl, pH 5.5) for 45 minutes and then rinsed. A large strip of 3MM filter paper (Whatman) was soaked in 20 X SSC (3M NaCl, 0.3M $Na_3C_6H_50_7.2H_2O$, pH7.0) and placed on a glass plate so that the ends of the paper were in a reservoir of 20x SSC, forming a wick. The gel was placed on top of the wet filter paper, then a correctly sized piece of nylon membrane (Hybond-N, Amersham), was placed directly onto the gel. Two pieces of 3MM blotting paper, pre-soaked in 20x SSC, were placed on top of the membrane. Air bubbles were removed carefully. Any exposed wick was screened off with Saran wrap (Dow Chemical Company), then a weighted stack of paper towels was placed on top.

Gels were blotted for were blotted for 5 hours to overnight. After blotting, the membranes were left to air dry, and then DNA was bound to the filter by exposing it to 1200μJoules of UV irradiation in a Stratalinker (Stratagene) or baked at 80°C for 1hour. Nylon membranes were stored in Saran wrap or between two sheets of 3MM paper at room temperature.

2.8.2 Transfer of bacterial colonies and phage plaques to filters and replication of filters

Plates containing phage plaques had to be placed at 4°C for at least 15 minutes prior to lifts being taken to ensure that the top agar was not removed during the procedure. Nitrocellulose (Schleicher and Schuell) or nylon (Hybond-N, Amersham) circular filters were marked with an asymmetric pattern of dots and placed carefully on the surface of the agar plate, avoiding air bubbles. The filter was left on the surface for approximately 10 secs while the pattern of orientation dots from the filter was transferred onto the bottom of the Petri dish. The filter was then removed carefully, avoiding smearing the colonies/plaques, and placed face up. The bacterial colonies on the agar plates were left at 37°C for several hours to 'recover'. A second lift could be taken from the phage plates if required after a further 15 minutes at 4°C.

2.8.3 Fixation of the bacterial colonies / phage plaques to filters by lysis

One sheet of 3MM paper was soaked in 10% SDS and the filters were placed face up and left for 5 minutes. The filters were then removed and placed on 3MM paper which had been soaked in denature solution (1.5M NaCl, 0.5M NaOH) and again left for 5 minutes before being transferred to neutraliser solution (1.5M NaCl, 0.5M Tris.HCl (pH 7.5))again on 3MM paper for 5 minutes. The final stages was to soak the filters in 2X SSC for 5-10 minutes. The filters were left to air dry and were then baked at 80°C in a vacuum for 20 mins (nitrocellulose) or UV irradiated as in section (nylon).

2.9 Radiolabelling of DNA

2.9.1 Preparation of DNA for probes

Probe DNA was in two forms, either in solution or contained within a gel slice. 25-50ng was labelled for a single hybridisation. DNA for gel slice hybridisation probes was excised from an Low Melting Point (LMP) gel after electrophoresis and diluted with 1-3 times (depending on concentration) its volume of water. The gel slice was melted at 65°C and vortexed, before storage at -20°C. Before use in a labelling reaction, the gel slice was remelted at 70°C.

2.9.2 Random Priming of DNA probes

This method is adapted from Feinberg and Vogelstein (1983 and 1984).

A labelling reaction with [α-^{32}P]-dCTP involves random priming from hexanucleotides and then polymerisation along the DNA strand catalysed by the Klenow fragment of *E.Coli* polymerase 1. A radiolabelled base is incorporated at every C nucleotide.

The DNA strands were firstly denatured by heating to 100°C for 5-10 mins. DNA in solution was then kept on ice to prevent reannealing of the strands. DNA in a gel slice form was allowed to cool slightly before the labelling reaction. The labelling reaction was carried out using either a Random Prime kit (Boehringer Mannheim) or a Hi-Prime kit (Boehringer Mannheim). For the random prime kit, 11µl total of DNA plus water was mixed with 3µl 10µCi/µl [α-^{32}P]-dCTP (Amersham), 1µl (2 units) Klenow enzyme, 1µl each of dATP, dTTP and dGTP and 2µl reaction buffer. For the Hi-Prime kit, 13µl of DNA plus water was mixed with 3µl 10µCi/µl [α-^{32}P]-dCTP and 4µl Hi-Prime (which contains the enzyme, nucleotides and buffer). The reactions were then incubated at 37°C for 1 hour (Random Prime or Hi-prime with a gel slice) or 10 mins (Hi-Prime with DNA in solution).

The percentage incorporation of the radiolabelled nucleotide was checked by TCA precipitation of ~0.5µl of the reaction mix on a GF/B circular filter (Whatman). If the incorporation was 50% or above, the unincorporated nucleotides were removed by running the probe through a Sephadex G-50 Nick column (Pharmacia Biotech). The storage buffer was removed and then the column was washed by running through ~500µl TE. The probe (~20µl) was then added to the top of the column, followed by 380µl TE. The probe was then eluted with a further 400µl TE, and collected in a tube containing 500µg sonicated salmon sperm DNA (Sigma). The probe was denatured by heating to 100°C for 10 mins and added to the prehybridisation solution in the bottle.

Hybridisation probes were "stripped" in order to minimise hybridisation occurring due to high copy number repetitive elements in the probe. The probe was denatured at 100°C for 10 mins in the presence of 500µg sonicated salmon sperm DNA and

250μg sonicated total human genomic DNA. It was then left to reanneal for 40 minutes at 68°C before addition to the hybridisation bottle.

2.10 Hybridisation Protocols

2.10.1 Hybridisation Solutions

(Pre)hybridisation solution

5x SSC, 10% dextran sulphate, 0.1% sodium pyrophosphate, 0.1% SDS and 5x Denhardt's solution. 5x Denhardt's solution contains 0.1% ficoll, 0.1% PVP and 0.1% BSA (Sigma). Filter before use.

2.10.2 Prehybridisation of filters

Under a solution of 2x SSC, the filter was placed between two slightly larger sheets of gauze. For hybridisation to multiple filters with a single probe, up to 3 layers of filters could be placed in the same bottle, with layers of gauze separating them. Air bubbles trapped between the filter and the gauze were removed. The filter and gauzes were then rolled up together, transferred to a glass hybridisation bottle (Hybaid) and unrolled onto the surface of the bottle.

Denatured sonicated salmon sperm DNA was added to the prehybridisation solution at a concentration of 100μg/ml. 12ml of prehybridisation solution was added to a small bottle and 18ml to a large bottle, and the bottle rotated in an hybridisation oven (Hybaid) at 68°C for a minimum of 1 hour before adding the probe.

2.10.3 Hybridisation conditions

For random-primed probes, hybridisations were carried out at 68°C overnight.

2.10.4 Washing conditions

Following hybridisation, the filters were removed from the bottle and separated from the layers of gauze. In a plastic tray, the filters were washed with 3 changes of ~500ml washing solution (0.1-2X SSC, 0.1% SDS) using gentle agitation. The temperature of the wash and the concentration of SSC depended on the washing stringency required. For random-primed probes, a high stringency wash (3x 15 mins in 0.1x SSC, 0.1% SDS at 68°C) was commonly used.

The filters were then wrapped in Saran wrap (Dow Chemical Company), avoiding creases.

2.10.5 Detection of hybridisation signal

Autoradiography :

The filters were placed in a light-tight cassette with a signal enhancing screen. They were then exposed to X-OMAT x-ray film (Kodak) for a length of time dependent on the amount of radiolabelled probe left bound to the filter (several minutes to several days). Filters hybridised to ^{32}P-labelled probes were exposed at -70°C. Stratagene Glogos II luminescent markers were used for alignment. The film was developed on an automatic x-ray film processor RGII (Fuji).

Phosphorimaging Alternatively, the filters were exposed to a phosphor screen (Molecular Dynamics) for hours to several days. The screen was then scanned on a PhosphorImager (Molecular Dynamics), where a laser beam converts the radioactive signal into a digital image, with variations in the pixel value proportional to the amount of radioactive signal present. The grey-scale image was adjusted as desired and was then printed on a laser printer.

2.10.6 Removal of radioactive probe from filters

Since hybridisation filters can be used several times with different hybridisation probes, removal of radioactive probes from the filters is often required. This can be achieved in one of two ways. Firstly, the filters could be washed in 0.4M NaOH at 50°C for 30 mins, to denature the probe DNA strand from the surface of the filter. The neutral pH was then restored by a 30 min wash in 0.1x SSC, 0.1% SDS, 0.2M Tris.HCl, pH 7.5, also at 50°C. Alternatively the filters could be subjected to one very stringent wash, in 0.1% SDS or in distilled water at 100°C for 5 mins.

The filters were then exposed to x-ray film overnight to check that all the probe had been removed.

2.11 Enzymatic Manipulation of DNA

2.11.1 Restriction enzyme digestion of DNA

Digestions of DNA with restriction endonucleases were carried out in the appropriate buffer at the temperature recommended by the manufacturer. Restriction enzymes were supplied by Boehringer Mannheim, NEB and Gibco BRL. 1-10μg of DNA was digested in 10-40μl containing 1/20 volume 0.1M spermidine (Sigma), using 1 unit of enzyme per μg of DNA cut. The reaction was left at the appropriate temperature for 1½ hours to overnight. Somatic cell hybrid, total human genomic and YAC DNA was always digested overnight. If two different enzymes were used, both of which required the same buffer, the digests were carried out simultaneously. Otherwise, after digestion with one enzyme, the sample was drop dialysed for 30 mins or ethanol precipitated. The appropriate buffer was then added and the second digestion carried out. When necessary, reactions were terminated by heating to 68°C or 80ºC for 15 minutes, according to the heat sensitivity of the enzyme (NEB Catalogue, 1992). Restriction digests which were to be run on gels were terminated by the addition of 1/10th of their volume of "stop mix" (100mM EDTA, pH8.0; 20% Ficoll and Orange G).

2.11.2 Dephosphorylation of 5' termini

Calf intestinal phosphatase (CIP) was used to dephosphorylate the 5' ends of the vector molecules before cloning. This prevents recircularisation of vector molecules during the ligation step.

7μg DNA was dephosphorylated with 0.1 unit CIP (Boehringer Mannheim) in 50μl 1x CIP buffer (10mM Tris.HCl, pH 8.3, 1mM $ZnCl_2$, 1mM $MgCl_2$) at 37°C for 30 mins. The reaction was stopped by the addition of 1μl of 0.5M EDTA. The dephoshorylated DNA was then chloroform/isoamyl alcohol then ether extracted and ethanol precipitated. The pellet was resuspended in 10μl TE.

2.11.3 Phosphorylation of 5' termini

Addition of a phosphate group to a DNA strand ('kinasing') is by transfer of the terminal phosphate of γATP. It is necessary to kinase oligonucleotides before they can be ligated to restriction enzyme digested DNA.

4μg oligonucleotide DNA was mixed with 1μl (10 units) polynucleotide kinase (PNK; Boehringer Mannheim) and 0.5-1mM ATP (Boehringer Mannheim) in 1/10 volume 10x PNK buffer (50mM Tris.HCl, pH 7.5, 10mM $MgCl_2$, 5mM DTT) in a total volume of 50μl. The reaction was incubated at 37°C for 30 mins. The sample was then chloroform/isoamylalcohol then ether extracted and ethanol precipitated. A recovery of 25% was assumed.

2.11.4 Ligation of cohesive termini

The insert DNA and the vector DNA were mixed on ice, typically in a ratio of 5:1 vector to insert. 10-100ng of vector DNA was normally used. The reaction was carried out in 10-50μl using 0.1 unit of bacteriophage T4 DNA ligase (Boehringer Mannheim). in 1x ligase buffer (66mM Tris.HCl, pH 7.5, 5mM $MgCl_2$, 1mM DTT, 1mM ATP). The ligation mixture was then incubated at 16°C overnight. The enzyme was then inactivated by heating to 65°C for 15 mins.

For a trial ligation, 100ng of bacteriophage λ digested with Hind III was religated under the same conditions. The extent of ligation could be visualised by running the religated product on a 1% agarose gel.

2.12 Oligonucleotides

2.12.1 Oligonucleotide Synthesis

Oligonucleotides were either synthesised (by Agnes Gallagher) as ammonium stocks on an Applied Biosystems 381A oligonucleotide synthesiser or were purchased from Genseis. Oligonucleotides for magnetic bead capture were synthesised incorporating a 5' biotin moiety. Oligonucleotides for use in the allelic association study were fluorescently labelled.

Oligonucleotides were precipitated from ammonium stocks by ethanol precipitation of 350μl of the stock. The precipitated DNA was resuspended in 500μl water and the concentration assessed by measuring the Abs_{260}. An absorbance of 1 corresponds to a concentration of 25μg/ml for single stranded DNA.
The oligonucleotides that were obtained from Genesis Ltd were in lyophilised form.

2.12.2 Oligonucleotide primer Design

The recombinant DNA inserts in the plasmid pBS were amplified using two published primers; 291 and 292 ('forward' and 'reverse' primers from Stratagene catalogue, see Table 2.2 for sequences).

Other primers were designed using the programme Oligo4 (Hybaid). Primers were between 18 and 46 nucleotides in length depending on the DNA source on which they would be used. They were chosen to be stable oligonucleotides with a T_m between 50°C and 70°C. A maximum hairpin structure or potential dimerism of 3 nucleotides (preferably less) was permissible, but only a 2 nucleotide internal match if it involved the 3' end. If possible the A+T : C+G ratio was near 50%.

In one instance, it was necessary to synthesise slightly degenerate primers due to unresolved sequence ambiguities.

2.12.3 Duplexing oligonucleotides

In certain cases it was necessary to form a duplex between two complementary (or partially complementary) oligonucleotides. The two oligonucleotides (150µg/ml each) were mixed together in 10µl 1x TM (10mM Tris.HCl, pH 7.5, 5mM $MgCl_2$) and the tubes were placed in a small volume of water at 90°C. The water was then left to cool to below 30°C over a period of 15 minutes to 1 hour.

2.13 Amplification of DNA by the Polymerase Chain Reaction (PCR)

PCR is a technique used for the amplification of a specific DNA sequence (Saiki et al, 1988; Mullis and Faloona, 1987). The specificity is provided by oligonucleotide primers complementary to the 5' ends of the two strands of the sequence to be amplified. These primers anneal to the template DNA strand and direct amplification by a thermostable polymerase. The PCR reaction is a series of cycles, each consisting of three steps. The first is a short high temperature step (92-94°C) to denature the template. This is followed by a reduction in temperature to allow the primers to anneal to the template (~50-65°C depending on the primers used), then finally the temperature is increased to 72°C (or 68°C for long PCR) to permit

extension from the primers by the polymerase. Multiple cycles of these three steps result in exponential amplification of the desired sequence.

2.13.1 PCR conditions

Commonly 100-200ng of template genomic DNA was amplified in 50µl in 0.5ml microcentrifuge tubes (Robbins Scientific), with 0.2µl (1 unit) Amplitaq (Perkin Elmer Cetus) in 1x reaction buffer (10mM Tris.HCl, pH 8.3, 50mM KCl, 1.5mM $MgCl_2$, 0.01% (w/v) gelatin) and 0.1% Triton X-100 (Sigma). The reaction employed 300ng of each primer and 200µM each dNTP (Advanced Biotechnologies). The reaction was overlaid with 40µl mineral oil (Sigma).

PCR programmes were run on an Hybaid Omnigene machine or Perkin Elmer Cetus DNA thermal cycler. The basic programme was 30 cycles of denaturation annealing and extension steps

The denaturation step was usually short and at a temperature of 93-94°C. Usually 1-2mins in the first cycle and 15-30 secs thereafter. The annealing step temperature is critical in determining the success of the PCR reaction, and is dependent on the structure of the primer. Primer T_m's were determined using the Oligo 4 programme (Hybaid) and annealing temperatures of ~5°C lower than the T_m was used. Annealing steps were 30 secs to 1 min long.

The extension step is carried out at 72°C. The extension time is also critical in determining the success of the PCR, since longer products need longer extension times (approximately 1 min per kb of sequence). The length of the extension time was generally the time estimated for effective elongation of the fragment required with a longer final extension time to ensure fragments were complete.

After the reaction, 10µl of the PCR reaction was visualised by electrophoresis on 0.8-2% agarose gels.

'Hot Start' PCR (Chou et al, 1992) was commonly used if increased specificity in the PCR reaction was required. The PCR reaction was set up as 2 separate mixes, one containing the DNA and primers, the other containing the enzyme, buffer and nucleotides. The mixes were heated to 90°C for several minutes, before combining them. Addition of the enzyme to the DNA at a temperature at which the DNA is

denatured helps to prevent mispriming before the initial denaturing temperature is reached.

'Touch Down' PCR (Don et al, 1991) was also used to increase specificity of the PCR reaction and involves decreasing the annealing temperature by 2^0C per cycle until the final annealing temperature is reached. This means that during the first few critical cycles, only highly specific primer-template interactions can occur, which selects for amplification of the correct product during the rest of the reaction. An annealing temperature 10°C above the calculated value was used in the first cycle, which decreased by 2°C each cycle until the final annealing temperature was reached.

2.13.2 PCR amplification of vector inserts from bacterial colonies

This method was adapted from Taylor (1991).

A toothpick stab of a bacterial colony was transferred to a tube containing 50µl of water. This tube was then heated to 96°C for 4 mins. 1µl of this was then used as the template for a PCR reaction. Primers used were Universal primers designed to the polylinker sequence of pBS, 291 and 292 (see table 2.2).

The PCR reaction mix was exactly as in section 2.13.1

The PCR programme used was:

Denaturing: 93°C for 30 secs in first cycle and 15 secs thereafter.

Annealing: 55°C for 30secs.

Extension: 72°C, 10 cycles of 1 min, 10 cycles of 2 mins, 10 cycles of 3 mins.

2.14 Sequencing of DNA

2.14.1 Cycle sequencing of DNA

Sequencing of DNA was done on an ABI automated sequencing machine (Applied Biosystems). Thermal cycling of sequencing reactions increases signal intensity and decreases sensitivity to reaction conditions. Dye terminator chemistry was used, that is dideoxy-nucleotides terminated the sequencing reaction at each base to generate the ladder of fragments. Each of the four dideoxy-nucleotides is labelled with a different dye, such that the attached computer can recognise the peak of dye corresponding to a particular base at each position.

Plasmid DNA for sequencing was prepared by alkaline lysis followed by PEG precipitation according to the protocols in sections 2.2.5 and 2.2.6. All other DNA was also prepared in this way where possible or was sequenced using this method from PCR products (see next section). The sequencing reactions were performed using a Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems) according to the manufacturer's instructions. Ideally, approximately 200ng plasmid DNA was used in the reaction, although results were produced with as little as ~20ng. 10ng of either primer 291 or 292 was used per reaction, depending on the orientation of the sequence required.

The cycle sequencing reaction was performed on a Perkin Elmer Cetus DNA thermal cycler. The programme was as follows: Place tubes in rack preheated to 96°C. Perform 25 cycles of 96°C for 30 secs, 50°C for 15 secs and 60°C for 4 mins. Rapid thermal ramp to 4°C and hold.

As much as possible of the mineral oil was removed with a pipette, then the product was ethanol precipitated, using Applied Biosytems NaAc and 3 volumes of cold 95% ethanol. The tubes were then placed on ice for 15-30mins, before centrifugation for 15-20 mins at 11,000rpm. The supernatant was then removed, and the pellet washed with 70% ethanol and spun at 11,000rpm for a further 10 minutes. The pellet was then vacuum dried for 3 minutes. Immediately before loading onto a sequencing gel, each pellet was resuspended in 4µl stop mix (1µl of 50mM EDTA, pH 8, 30mg/ml Blue Dextran to 5µl deionised formamide) and heated to 90°C for 2 minutes to allow the DNA to denature.

The samples were loaded onto a 6% acrylamide gel (prepared by Agnes Gallagher), which was run in 1x TBE at 2500V, 45mA for 12 hours. The sequence data was gathered by the computer and displayed as a visual output of differently coloured peaks, or DNA sequence which was directly amenable to computer analysis.

2.14.2 Sequencing of PCR products

This method has been modified from Winship, 1989.

The PCR product was first purified from a gel in order to separate it from primer dimers and any spurious products. A large amount of DNA is required in the reaction itself, so in some cases the product of several PCR reactions was pooled and concentrated into a smaller volume by ethanol precipitation. The resulting product was then gel purified by running it on a 1% (typically) LMP agarose gel and excising the required band. The DNA was recovered from the gel slice by agarase treatment and ethanol precipitation. The precipitate was resuspended in 10-40μl and the DNA concentration checked by running 1μl on a gel next to a known amount of size marker. Approximately 200ng DNA was used per sequencing reaction.

Alternatively, the PCR product was purified by excising the band from a 1% agarose gel. This gel slice was then placed at the top end of a 200μl filter tip (Rainin Instrument Company Inc.) placed in an epindorff tube and spun at 11,000rpm for 15 minutes at room temperature. The resulting solution was then drop dialysed as in section 2.7.3 for 45 minutes. A small aliquot of the product was then run on a gel in order to determine the concentration of the DNA and up to 1μg was used in the sequencing reaction depending on the size of the product.

Reactions were stored at -20°C until loading on a gel. Prior to loading they were heated to 80°C for 2 mins to denature the DNA.

2.14.3 Analysis of sequence data

The sequences obtained were edited and manipulated using GCG (Genetics Computer Group) programmes available through the HGMP Computing service (Devereux et al, 1984). Database searches were performed using the GCG Blast (Altschul et al, 1990) and Fasta (Pearson and Lipman, 1988) programmes.

2.15 Coincident Sequence Cloning

Coincident Sequence Cloning (CSC) is a method for isolating the sequences shared between two DNA resources (see section 6.1). This is achieved by denaturing the two DNA sources, mixing them ('integration') and then recovering the duplexes which are formed between the two resources (inter-resource duplexes, IRDs). In

this context, the two resources used were a pooled resource of PAC DJ4B9/3 and cosmids B01519 and I0142 from chromosome 1 and a pooled cDNA resource consisting of three ages of foetal brain cDNA, so that the product should be transcribed sequences from the PAC and cosmids.

At all times during the preparation of the Source DNAs, their integration and cloning, great care was taken so that the samples would not become contaminated. All materials used were dedicated solely to CSC, fresh solutions were made up and autoclaved, and all labware was washed with water and then ethanol or autoclaved where possible.

2.15.1 Preparation and digestion of the Genomic DNA resource

PAC and cosmid DNA was isolated using the same method as described for large scale preparation of plasmid DNA (i.e. CsCl equilibrium density gradient; see section 2.1.7). The PAC and cosmid DNA were then run out on a 1% LMP gel to assess the concentration of the DNA. The PAC and cosmid DNA was pooled in the ration of 4:1:1 . 50ng of the pooled DNA was ethanol precipitated and air dried.
The resulting pellet was resuspended in 5µl 1x restriction enzyme buffer A (Boehringer Mannheim) and digested with Sau3AI at 37°C for 1 hr. The enzyme was heat killed by heating at 70°C for 15 mins. The sample was then placed on ice.

2.15.2 Preparation and digestion of the cDNA resource

Oligo-dT primed poly A selected cDNA was prepared from three ages of foetal brain tissue using an Invitrogen *Fasttrack* mRNA isolation kit according to manufacturers instructions. Briefly the technique involved mechanical homogenization of the frozen tissue in lysis buffer which also inactivates any RNases released as the cells lyses. After homogenisation the cell lysate was left at 45°C for 15-60 minutes to allow the cells to fully lyse and complete digestion of proteins and ribonucleases to occur. 950µl of 5M NaCl was added and mixed by inversion. Any remaining DNA was sheared by passing the lysate 3-4 times through a sterile plastic syringe fitted with a 19 gauge needle. Oligo dT cellulose was then added to the lysate and the mixture incubated for 15-60 minutes to allow the mRNA to bind to the cellulose. (A typical mammalian cell contains about 10-11 g of RNA of which only 1-5% is poly A RNA. The rest is made up of rRNA (80-85%) and tRNA (15-20%). To separate the

heterogeneous population of mRNA from the other RNA in the cell, affinity binding to oligo dT cellulose is used exploiting the major characteristic of mRNA, polyadenylation, to obtain pure intact mRNA).

Once the mRNA has bound the oligo dT cellulose is pelleted by centrifugation at 3000Xg and the supernatant discarded. The cellulose then undergoes a series of washing stages to remove unwanted cell debris before elution of the mRNA from the cellulose oligo dT resin. The mRNA is then precipitated with 0.15 volumes of 2M sodium acetate and 2.5 volumes of ethanol by freezing on dry ice until solid and then centrifuging at 16000Xg for 15 minutes at 4°C. The pelleted mRNA then used for cDNA synthesis.

The Invitrogen *copykit* cDNA synthesis kit was used to convert single stranded mRNA into double stranded cDNA (based on a modified version of Gubler-Hoffman (1983) method of cDNA synthesis) according to manufacturers instructions. Briefly this involved denaturing the RNA at 65°C for 10 minutes and then in rapid succession adding placental RNase Inhibitor, reverse transcriptase buffer, 100mM dNTP 80mM sodium pyrophosphate and reverse transcriptase and incubating at 42°C for 60 minutes to allow the first strand cDNA to be synthesised (see manufacturers handbook for further details) this reaction was then put on ice before second strand cDNA synthesis was carried out by the addition of second strand buffer, 1mg/ml BSA, 10mM β-NAD, 1M DTT, RNase H, *E.coli* DNA ligase and DNA polymerase. This was then incubated at 15°C for 90 minutes and then a further 30 minutes at room temperature. The reaction was then stopped by heating to 70°C for 15 minutes prior to the addition of T4 DNA polymerase. This was incubated at room temperature for 10 minutes (to blunt end the cDNA) and 2µl of 0.5M EDTA was added to stop the reaction. The cDNA was then cleaned by phenol/chloroform precipitation and then the addition of 4M ammonium acetate and 3 volumes of ethanol. After 15 minutes on dry ice this was spun at 3,000Xg for 15 minutes at 4°C and the pellet retained and dried.

The three ages of foetal cDNA used were 15.3 weeks, 12.2 weeks and 10.5 weeks of gestation. These cDNA were separately extracted and an aliquot run out on a 1% agarose gel in order to determine their concentration. The cDNA s were then pooled in equal molar proportions. 100ng of this pooled resource was then ethanol precipitated and air dried. The pellet was resusupended in 5µl 1x restriction enzyme buffer A (Boehringer Mannheim) and digested with *Sau3AI* at 37°C for 1 hr. The

enzyme was heat killed by heating at 70°C for 15 minutes. The sample was then placed on ice.

2.16 'Catch Linkering' of Genomic and cDNA resources for CSC

Catch linkering is a method for amplifying DNA, for instance a whole PAC, into a simplified resource. It is an alternative to Alu-PCR (Nelson et al, 1989), yet is not dependent on the presence of any specific sequence elements other than frequent restriction enzyme sites. DNA is digested with Sau3AI, and then a short double stranded 'catch linker' is ligated onto each cohesive end. The intervening DNA can then be amplified by PCR using primers designed to the catch linkers. Catch linkered material can be used as a substrate for Coincident Sequence Cloning.

2.16.1 Preparation of catch linkers

Catch linker oligonucleotides were 477 / 479 for genomic DNA resource and 727 / 731 for cDNA resource (see Table 2.2 for sequences of these primers).

Oligonucleotide 477 was synthesised with and without biotin at the 5' end. Oligonucleotides 479 and 731 were phosphorylated at the 5' end to enable them to be ligated to the restriction enzyme digested DNA. 1μg of each oligonucleotide was duplexed in 10μl TM (10mM Tris.HCl; 5mM $MgCl_2$ pH 7.4) to form the double stranded linker.

2.16.2 Ligation of catch linkers

For the genomic resource, 200ng linker DNA was ligated onto ~50ng digested genomic or cDNA, in 20μl volume at 16°C overnight. For the cDNA resource 500ng of catch linker was ligated to 100ng cDNA in a 50μl volume at 16°C overnight. The sample was then briefly spun and placed on ice.

2.16.3 PCR amplification

1μl of the ligated material was then amplified by PCR, using 600ng oligonucleotide 477 or 727 as the primer. Biotinylated 479 was used if labelled product was desired (see Table 2.2).

The PCR programme used was:

Denaturation: 94°C for 15 secs

Annealing: 54°C for 30 secs

Extension: 72°C for 1 min (cycles 1-10), 2 mins (11-20) and 4 mins (21-35).
10µl of the PCR product was examined on a 1% agarose gel. The remainder was ethanol precipitated and resuspended as appropriate to future use.

2.17 Preparation of Genomic DNA

The genomic resource was catch linkered (section 2.13) using biotinylated 477 as the PCR primer.
Before use in the CSC experiments it was necessary to purify the catch linkered product from a gel, to remove primer dimers. It is preferable to do this without staining the DNA with ethidium bromide or exposing it to UV irradiation.

The catch linker PCR product was ethanol precipitated and resuspended in 10µl dH_20. 9µl of the sample was then run on a 1.5% LMP agarose gel in a midi-gel apparatus, without ethidium bromide staining. The remaining 1µl was run in a parallel, test track next to a ϕX HaeIII marker track. When the loading dye had reached the bottom of the gel, this test and marker tracks was sliced away from the gel and stained with ethidium bromide. The DNA in these tracks was then viewed on a UV transilluminator. The top and bottom positions of the catch linker product with respect to the marker were noted, and the gel was nicked where it should be cut to remove the primer dimers. This test track was then aligned alongside the rest of the gel, and the gel sliced across at the position of the nick.

In order to concentrate the product DNA into a smaller volume of gel, the orientation of the gel was reversed and the product run back up towards the wells. Loading dye placed in a hole at the bottom of the gel before turning it round was used to judge when the run was complete. A second ϕX HaeIII marker track was sliced off and stained, and nicks made at the positions corresponding to the top and bottom of the catch linker product smear. This track was then aligned with the track containing the product, and the product was excised from the gel, guided by the nicks in the marker track.

The DNA was purified from the gel slice by agarase treatment and then was ethanol precipitated overnight. The resulting pellet was resuspended in 5µl dH_20. The

concentration was estimated by running 0.5µl on a 1% agarose gel. The remaining 4.5µl (~800ng) was used in the HF-CSC experiment.

2.17.2 Preparation of cDNA

20µg cDNA which had been catch linkered using 727 as the PCR primer was digested to completion with Sau3A restriction enzyme. The Sau3A digestion is never entirely complete with 1-5% of the linker sequences being retained (most likely due to some of the cDNA being single stranded and therefore impossible to restriction enzyme digest). This allows both the end ligation CSC method to be performed using primers specific to the capture oligonucleotide sequences (789/596) and hybrid fishing CSC to be carried out using primer (727) designed to the catch linkers which have not been entirely removed from the cDNA resource.

2.17.3 Preparation of 'blocking' DNA

Human Cot1 DNA was used to pre-anneal to the denatured genomic and cDNA resources, to prevent inter-resource hybridisation due to high copy repeat elements. Two 10µg aliquots of Cot1 human genomic DNA were ethanol precipitated and vacuum dried.

2.17.4 Integration of the YAC and cDNA

200ng genomic DNA and 10µg cDNA were made up separately to 8.4µl in water. The 10µg Cot 1 DNA which was ethanol precipitated was resuspended in this 8.4µl for both the cDNA and genomic resource. 1.8µl 1M NaOH was added to each of these resources (with an appropriate time delay) and the tubes placed at 37°C for 5 mins to ensure the denaturation is complete. The tubes were then warmed to 50°C and 9.8µl FNET/HCl (6µl formamide, 2µl NET (400mM Tris. HCl, pH7.8, 2.5M NaCl, 50mM EDTA), 1.8µl 1M HCl) at 50°C was added to each to neutralise the alkali. The neutral pH of each was confirmed by spotting 0.5µl of each onto Indicator paper. The tubes were then placed at 45°C for 0.5 hours (genomic resource) and 3 hours(cDNA resource) for pre-annealing with the blocking DNA. After this, the two DNA resources were combined and left submerged at 45°C overnight to permit the formation of the IRDs.

The integration and subsequent steps were also performed using controls of 'cDNA only' and 'no DNA'. 25µl M280 streptavidin coated magnetic beads (Dynal) were washed in phosphate buffered saline/BSA according to the manufacturer's instructions and then incubated for 15 mins at RT in 100µl TEN-S/P (0.1% SDS, 0.5mg/ml PVP in TEN buffer (TEN buffer: 10mM Tris.HCl, pH7.4, 1mM EDTA, 100mM NaCl)). This buffer was then replaced with 100µl TEN-S/P containing 10µg sonicated salmon sperm DNA and incubated for a further 15 mins. This buffer was then removed and the beads resuspended in 25µl TEN-S/P.

The duplexed DNA was made to 0.1% SDS, 0.5mg/ml PVP and then this solution was used to replace the bead solution. The DNA/bead mixture was incubated for 30 mins at room temperature with occasional mixing.

Washes were performed with 6 changes of 150µl TEN at room temperature, then 8 changes of 150µl 0.1x SSC at 68°C. The beads were resuspended in 20µl of TEN and heated to 60°C before adding in 250ng of 732 and 735 (kinased) capture oligonucleotides (see Table2.2 for sequences). Then the solution was allowed to cool to room temperature over a 1hour period. The beads were washed with 4 changes of 100µl of TEN and the beads resuspended in 20µl Taq ligase buffer (Cambio). A 5 minute ligation was then performed at 45^0C using 10 units of Taq DNA Ligase. The reaction was stopped by adding 1µl of 500mM EDTA.
The beads were then further washed with 3 changes of 100µl of TEN at 45°C and 5 changes of 100µl 0.1X SSC at 65°C.
The products were then eluted from the beads by heating to 90°C for 4 mins in 10µl dH_20 before immobilisation of the beads and recovery of the buffer/eluate DNA. The eluate was then passed through a further magnetic separation to remove all traces of beads.

2.17.7 PCR amplification of the products

1µl of the eluate and controls were amplified in duplicate with 600ng of the cDNA linker primer 727 (HF-CSC) or 789 and 596 (EL-CSC) (Table 2.2). The programme used was:
Denaturation: 94°C for 30 secs in the 1st cycle and 15 secs thereafter.
Annealing: 54°C for 30 secs

Extension: 72°C for 1½ mins (cycles 1-10), 2½ mins (cycles 11-20), 4 mins (cycles 21-29) then 8 mins (cycle 30)

The PCR reactions were then visualised on a 2% agarose gel.

2.17.8 Cloning of the product DNA

A library was made from the end ligation product DNA by cloning the PCR product into pBS vector.

20µl of the product DNA was chloroform/isoamyl alcohol then ether extracted and ethanol precipitated. The pellet was resuspended in 5µl. The DNA was then heat shocked at 65°C for 5 mins and cooled to RT, to disentangle the molecules. The concentration was checked by running 1µl on a 2% agarose gel, and then the remainder was digested with EcoRI in a total of 10µl. The enzyme was heat killed at 70°C for 10 mins, and then the sample was ethanol precipitated. The pellet was resuspended in 10µl and subjected to heat shock as above.

20ng of the digested product DNA was ligated to 20ng pBS DNA (digested with EcoRI and phosphatased), in a total of 20µl at 16°C overnight. The ligation mix was then drop dialysed and 1µl was used for electro-transfomation of competent cells.

2.18 Screening a full length foetal brain library

The full length foetal brain library which was screened was obtained from Clontech.

Library details:

Human full length foetal brain 5'-STRETCH PLUS cDNA library

Titre: $>10^8$ pfu/ml

Vector: λgt10

Host Strain: C600 Hfl

Cloning site: *EcoRI*

Priming method: oligo dT and random primed

mRNA source: pooled from 7 Caucasian foetuses 20-26 weeks

Average insert size: 0.4-2.5Kb

The library was plated out according to manufacturers instructions onto L-agar plates. Briefly the library stock is diluted and combined with the host *E.coli* strain (C600) which has been grown in the presence of maltose and magnesium sulphate at 37°C until confluent. This solution, after 15 minutes of further incubation at 37°C with agitation, was mixed with LB soft top agar and poured on the surface of a plate which already had set L-agar in it. This was then incubated at 37°C overnight. The plates were then removed and placed a 4°C for at least 15 minutes to prevent the top agar from being removed during transfer of the plaques to membranes. The plaques were then transferred to nitrocellulose membrane filters and fixed as in section 2.8.2 and 2.8.3. These filters were then hybridised with probes derived from the CSC product library as in section 2.8. Once positive plaques were identified plugs were taken. This was achieved using a large 1ml tip to remove the region which had a positive plaque in it as assessed by lining up the autoradiogram and original agar plate. These plugs were then allowed to soak in lambda dilution buffer (10X dilution buffer: 58.3g 1M NaCl 24.65g 0.1M $MgSo_4.7H_2O$ 350ml 1M Tris.HCl (pH 7.5), 50ml 2% gelatin make up to 1L with water and autoclave) overnight and replated as above in order to obtain single positive plaques. Once a single plaque has been obtained in plug form this is again allowed to soak overnight in dilution buffer and then the solution is diluted 1/1000 and this is heated to 90°C for 5 minutes. 2µl of this solution can then be used in a PCR reaction with Primers 391 and 392 which are designed to the λgt10 vector sequence and amplify up the cDNA insert. Several PCR products can then be pooled together and used in sequencing reactions as described in 2.14.

Table 2.2 Sequences of oligonucleotides used for PCR amplification and DNA sequencing

Name	Oligonucleotide sequence 5'→ 3'	Function/comments
	Actinin PCR and sequencing	
J101	GCT TCT GTA ATC ACT CAT CCC	ACTN 2 PCR
J102	AGG ATA CTG GTT TCT GAC TTG	ACTN 2 PCR
J103	GAG ACT GAC ACG ACT GAG CAA	ACTN 3 PCR
J104	TGG CTG GCT TTT CTC TTA GGC	ACTN 3 PCR
	Plasmid vector primers	
291	CAG GAA ACA GCT ATG AC	pBS primer reverse
292	GTA AAA CGA CGG CCA GT	pBS primer forward
M13	TGT AAA ACG ACG GCC AGT	forward
M13	CAG GAA ACA GCT ATG ACC AT	reverse
	EST Mapping and sequencing	
*K290	GCA CAG AGA ATG AAT GAT AAG	Z39242 sequencing
K97	TGA CAG GGG TGA TTC TCC	Z39242 sequencing
K229	CAC CTG GAA GCA TCT GTT	Z39242 sequencing
*K872	CTT TTA GAT TGT GGG TAT TAT	Z39242 sequencing
K204	CTC AGA AAT CAC CAC TAA	Z39242 sequencing
K301	CTG CCA TAC TTG TTT TCC	Z39242 sequencing
K568	ATG GGA ATG GAA ATG GTG	Z39242 sequencing
K85	TAG CGG GTG TGA AGT GGT ATC	Z39242 sequencing
K289	ACT TCA TAA GTA TGC AAA ACA	Z39242 sequencing
	Primer pairs for EST PCR	
Z19327	AGG CAC AAG TTT AAG CAG CAT TGC CAG TCT GTA TTG AGG GGT	
Z28525	AGT CTC CAC AAA TGG TCT GCT AAT GCC GTA TGT TGT AGG GGT	
Z28611	GTC TGT CTC TGC GGC TTC TGT CCT TCA AGT TCC AAG CGT CT	
Z28653	TTG AAA ACA CTG AAA ACC TCT TGA ACT GAA CCT TGG ACT GAT	
F00203	AGT GGT CGC CAA GTA AGA GG CTT GGC ATG TGA GGG CTA TT	
F00755	GTG GAG CAC ATT ATA GAC AG AAT GAG TTG TTG TTG TTG AT	
F00845	GTG GTT ACA TCT ATT CTT GCT	

	TGC CTT GGA TAC TTG GCA TTT	
Z45606	GAA ATC TGT CCC GTA AAG ATA AGA CAC CAG CAG TAG	
Z42258	ATG TTC TTT TCA GCG GTC AAA TAG GGT TGT AAA AGG	
Z42924	TAA CCT ATT TCC ACA TTC TGT TAA AGA TGA CTG TGA	
Z46134	ATG TTT ACC TCT TTC TTT GAA AAG ATA CTC AAG GTT	
Allelic association study primers		
D1S1621	TTT CTC ACC TTT AAA TGT CAT CA	Forward
D1S1621	CCA GTA CGC AGA TGG TCC TA	Reverse
D11S931	AGT TTG GTA GGT ATT CT	Forward
D11S931	GAG AAA TAG TAT GTG TTT GCC	Reverse
Coincident Sequence Cloning		
477	CCG AAT TCT AGA GTC GAC C	+/- Biotin at 5' end genomic resource catch linker (lower strand)
479	GAT CGG TCG ACT CTA GAA TTC GG	genomic resource catch linker (upper strand)
727	GCG AAT TCT AGA CTG CAG G	cDNA resource catch linker (upper strand) & primer
731	AAT TCC TGC AGT CTA GAA TTC GC	cDNA resource catch linker (lower strand)
732	GGA CGG GTC GAC ACG CGA GGA CCG AAT TCT AGA GTC GAC C	End Ligation capture oligo 1 (part complement of 479)
735	GAT CGG TCG ACT CTA GAA TTC ACC CGT GCT ACC GGA ACG	End Ligation capture oligo 2(part complement of 477
596	GGA CGG GTC GAC ACG CGA GG	End Ligation product primer 1
789	CGT TCC GGT AGC ACG GG	End Ligation product primer 2
CSC clone sequencing primers		
M188	CCA GAA GTT TGT CAC CAG	CSC clone 3A6 sequencing primer
M187	TGA GGC AGG AGA ATC GCT	CSC clone 3A6 sequencing primer
M305	TTG TCT GCC TGA AAA TGT	CSC clone 3A6 sequencing primer

M306	CCA GGG ACA AAG TTW WCA	CSC clone 3A6 sequencing primer
N164	TAA CAA CTA ACC TCC CCT	CSC clone 3A6 sequencing primer
I184	CTT TCC ACC TTC TGC CAT	CSC clone 3A6 sequencing primer
N811	GCA GGG AAG CAG GAG AAA	CSC clone 11B6 internal primer
N812	TCC CAG AAA ACA ATA CCC	CSC clone 11B6 internal primer
Vector Primer for EST clone 299277		
T3	AAT TAA CCC TCA CTA AAG GG	Vector pT3T7
T7	CAT TAT GCT GAG TGA TAT CCC G	Vector pT3T7
Vector Primers for phage cDNA clones		
391	AGC AAG TTC AGC CTG GTT AAG T	Vector λGT10
392	TTA TGA GTA TTT CTT CCA GGG	Vector λGT10

Chapter Three

High Resolution Mapping of The Actinin Genes

ACTN 2 and ACTN 3

3.1 Introduction

α-actinin is part of the spectrin gene super family, a diverse group of cytoskeletal proteins, which also include the α and β-spectrins and dystrophin. These family members are characterised structurally by a central rod domain which is composed of 4 (α-actinin) to 24 (dystrophin) repeat units and EF-hand (helix-loop-helix motif) calcium binding domain that may or may not be functional. There is also a highly conserved amino terminal actin binding domain in spectrin, α-actinin and dystrophin and a carboxyl terminal domain which appears to be unique to dystrophin. The spectrins and dystrophins are thought to play a role in maintaining cellular integrity and flexibility as components of the membrane cytoskeleton. The α-actinins have more diverse functions but all serve as actin binding and crosslinking proteins in both muscle and non muscle cells. α-actinin is rod shaped and dimerizes in an anti-parallel fashion. In skeletal muscle, α-actinin is a major component of the Z-discs. In non muscle cells α-actinin is found along microfilament bundles where it may mediate membrane attachment at adherns-type junctions in a dynamic manner, regulated by the calcium binding EF hands. In smooth muscle α-actinin is found at dense bodies and dense plaques and has similar anchoring functions (Blanchard et al 1989, Matsudaira 1991, Beggs et al 1992).

A number of different isoforms of α-actinin have been identified by biochemical studies such as isoelectric focusing but many of these are thought to result from post translational modifications of the protein. Brain isoforms of α-actinin and spectrin exist, the latter being known as fodrin.

α-actinin genes have been described in the invertebrate Dictyostelium discoideum, Caenorhabditis elegans and Drosophila melanogaster and in the vertebrates chicken, mouse and rat. Three human α-actinin genes have been described. α-actinin 1(ACTN1) maps to chromosome 14q22-24 and is a non-muscle cytoskeletal isoform whereas ACTN 2 (1q42-q43) and ACTN 3 (11q13-q14) are expressed in skeletal muscle (Beggs et al 1992). Amino acid sequence comparison demonstrates that these genes are highly homologous to each other and are also conserved across species.

Defects in the actin crosslinking proteins have been associated with several human diseases. Mutations in the spectrin gene which result in a quantitative deficiency of spectrin affecting membrane binding properties and oligomerisation, can cause some forms of haemolytic anaemia. Becker and Duchenne muscular dystrophy are

genetically inherited wasting diseases of skeletal muscle which are caused by deletions in the muscular dystrophy gene resulting in truncation of the protein product, dystrophin.

α-actinin has not been associated with any human diseases as yet, although it was a candidate gene for Nemaline Myopathy which is a neuromuscular disorder characterised by the presence in skeletal muscle of nemaline rods composed, in part, of α-actinin. α-actinin has since been excluded as being involved in this disorder (Tahvanainen et al 1994). North and Beggs (1996) found a deficiency of ACTN 3 in merosin positive congenital muscular dystrophy (a subset of patients with congenital muscular dystrophy who are deficient for the extracellular matrix protein merosin). It is likely that this may be a marker of the condition and it is undetermined as to whether this deficiency reflects a gene mutation or is a secondary effect.

Other structural proteins which have been described as being involved in disease processes include myosin in which a defect in the myosin VIIA gene is responsible for Usher syndrome type 1B (Weil et al 1995). Usher syndrome is the most frequent cause of deaf-blindness in humans. Interestingly the gene responsible for this disorder (USH1B) maps to 11q13.5 and there has been a report of schizophrenia co-morbid with Usher syndrome (St Clair 1994). Another gene for Usher syndrome has also been mapped to chromosome 1q41. Abnormal expression of the microtubule associated proteins (MAPs) MAP2 and MAP5 in specific sub-fields of the hippocampus of schizophrenics has been reported and it has been suggested that this could underlie some of the cytoarchitectural abnormalities described in schizophrenia as well as possibly impairing signal transduction in the affected dendrites (Arnold et al 1991). Microtubule associated proteins have also been implicated in temporal lobe abnormalities in schizophrenics due to faulty assembly of these proteins resulting from a deficiency of trophic forms of non NMDA receptors (Kerwin 1993).

Brain α-actinin has been studied previously in chick brain to elucidate any possible involvement it may have in neurotransmission (Zheng and Babitch 1995). Chemical neurotransmitters are stored within the nerve terminals in synaptic vesicles that are often associated with cytoskeletal components and the presynaptic plasma membrane. When nerve stimulation causes influx of calcium, neurotransmitter is released and vesicles are recruited for release by rearrangements of the nerve ending cytoskeleton of which actin is a component. However, studies by Zheng and

Babitch (1995) indicated that chick brain α-actinin was localised in the cytoplasm and the level in nerve endings was minimal. They concluded that α-actinin was not likely to be involved in neurotransmission.

Recently Wyszynski et al (1997) identified α-actinin 2 as being a brain postsynaptic density protein that co-localises in dendritic spines with NMDA receptors and the putative NMDA receptor clustering molecule PSD-95 in the rat brain. NMDA receptors exist *in vivo* as heteromultimeric complexes consisting of the essential NR1 subunit which is co-assembled with various units of the NR2 subfamily which in turn are associated with PDS-95 synaptic proteins. α-actinin 2 binds to the cytoplasmic tail of both NR1 and NR2B by its central rod domain. NR1 α-actinin binding is directly antagonised by calcium / calmodulin. The authors speculate that calcium / calmodulin may displace α-actinin 2 from NR1 subunits, to which it may bind at resting intracellular calcium levels, in response to postsynaptic calcium influx. This mechanism may contribute to calcium dependent inactivation and run down of NMDA receptors and may also lead to receptor detachment from the actin cytoskeleton and their redistribution during synaptic activity. Hence α-actinin may play a role in both the localisation of NMDA receptors and their modulation by calcium.

α-actinin is also involved in the organisation of networks and bundles of actin filaments in neural growth cones (the growing tips of axons or dendrites). During development growth cones guide growing neurites to their appropriate target site and are the direct precursors to synaptic and neurosecretory terminals, undergoing dramatic structural changes upon target recognition (actin-based motility of lamellae ceases and specialised neurosecretory release sites are organised) (Forscher and Smith 1988). α-actinin may be involved in the organisation of actin at the growth cone margin and in its dynamic protrusions (Letourneau and Shattuck 1989). α-actinin may also be a linker protein between receptors for the extracellular matrix, integrins and actin filaments and be involved in the adhesion process of growth cones to the substratum (Sobue 1993). α-actinin-actin interactions may therefore cause bundling of actin filaments within filopodia or anchorage to the filopodia membrane and may induce retraction of growth cones or encourage reorganisation of actin microfilaments during growth cone movement. Hence, it is possible that abnormalities in α-actinin may have an effect on neuronal migration and navigation.

The α-actinin genes, ACTN 2 and ACTN 3, have been described by Beggs et al (1992) as mapping to chromosome 1q42-q43 and 11q13-q14 respectively. The translocation breakpoint described in the K26 pedigree which co-segregates with schizophrenia and associated disorders resides at t(1:11)(q43,q14.2). Primarily due to both the location of these two closely related genes on both the translocation chromosomes, in the vicinity of the breakpoint which might explain the mechanism of chromosomal rearrangement through illegitimate recombination between paralogs, and because of their biological functions, ACTN 2 and 3 were considered to be of interest as potential candidate genes. The aim of these experiments was to further determine the candidacy of these genes by first determining their precise location with respect to the translocation breakpoint by high resolution mapping, utilising available somatic cell hybrids and YACs from the contigs which had been produced around the breakpoint on chromosome 1 and 11.

3.2 High Resolution mapping of the ACTN 2 and ACTN 3 actinin genes

3.2.1Production of YAC and Somatic cell hybrid panels;

YAC contiguous clone maps have been assembled which extend over the region disrupted by the translocation on chromosomes 1 and 11(see section 1.16). These contigs span a 3Mb region in the case of chromosome 11 and 2Mb in the case of chromosome 1. Restriction enzyme mapping panels were constructed for both chromosome 1 and 11 using DNA from the YACs in these respective contigs (Figure 3.1 and 3.2). 5μg of YAC DNA was cut with restriction enzymes *EcoR1* or *HindIII* and then run out on a 1% agarose gel and Southern blotted using standard procedures (see section 2.8). These filters were then used in hybridisation reactions with probes made from the actinin genes, ACTN 2 and ACTN 3.

Several somatic cell hybrids were available containing fragments from human chromosome 11 including the translocation derived 11 chromosome, the human chromosome 11 only hybrid J1CI4 and the series of 4 WJX hybrids derived from J1CI4 by X-irradiation (Fletcher et al 1993). A derived 1 translocation chromosome carrying somatic cell hybrid was also available as was a human chromosome 1 only carrying hybrid A9. A full list of these hybrids is described in Table 3.1. These hybrids have been thoroughly characterised using a large number of chromosome 11 markers thus defining each interval on chromosome 11 (Evans et al 1995).

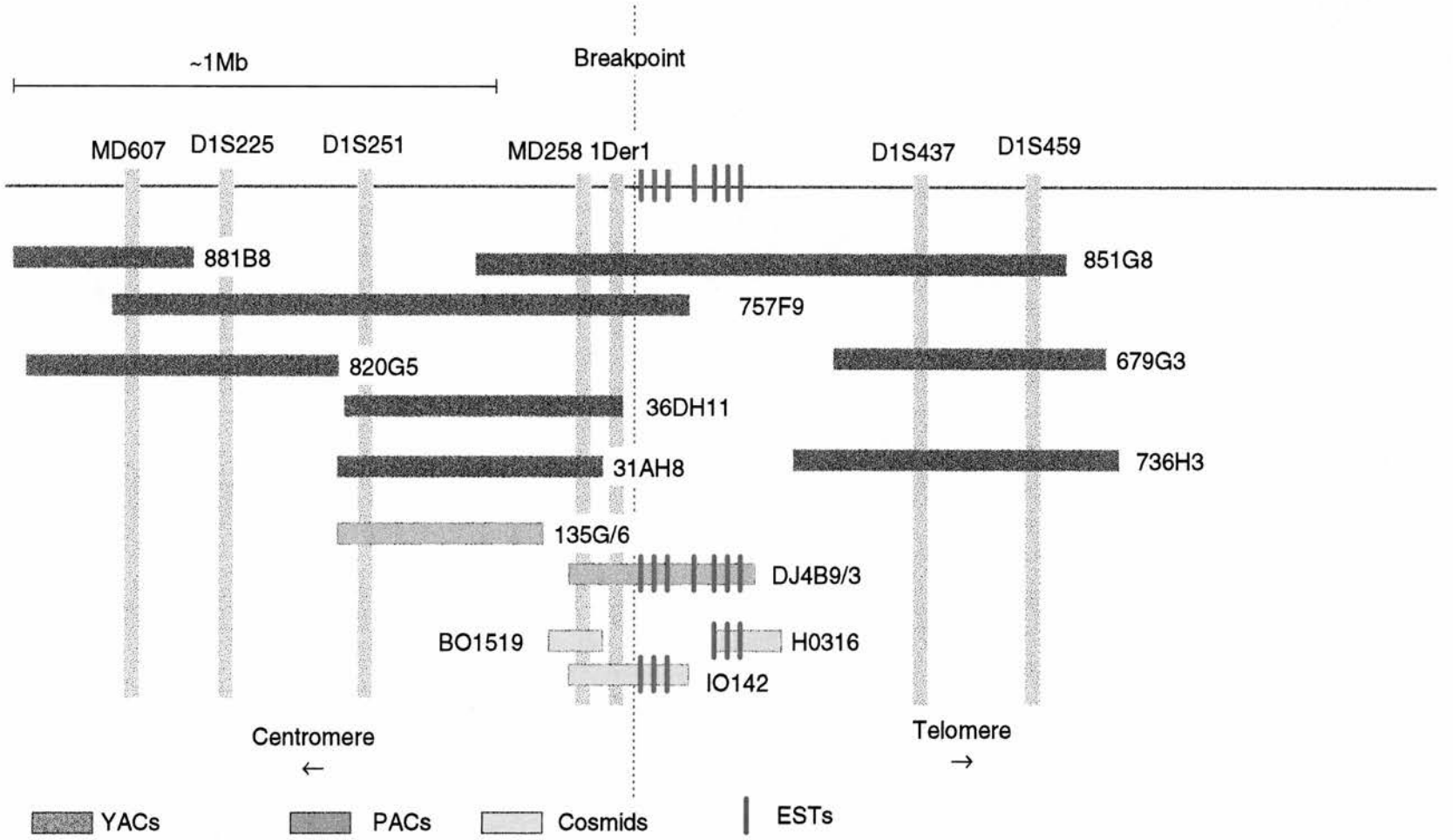


Figure 3.1 Chromosome 1 YAC contiguous clone map

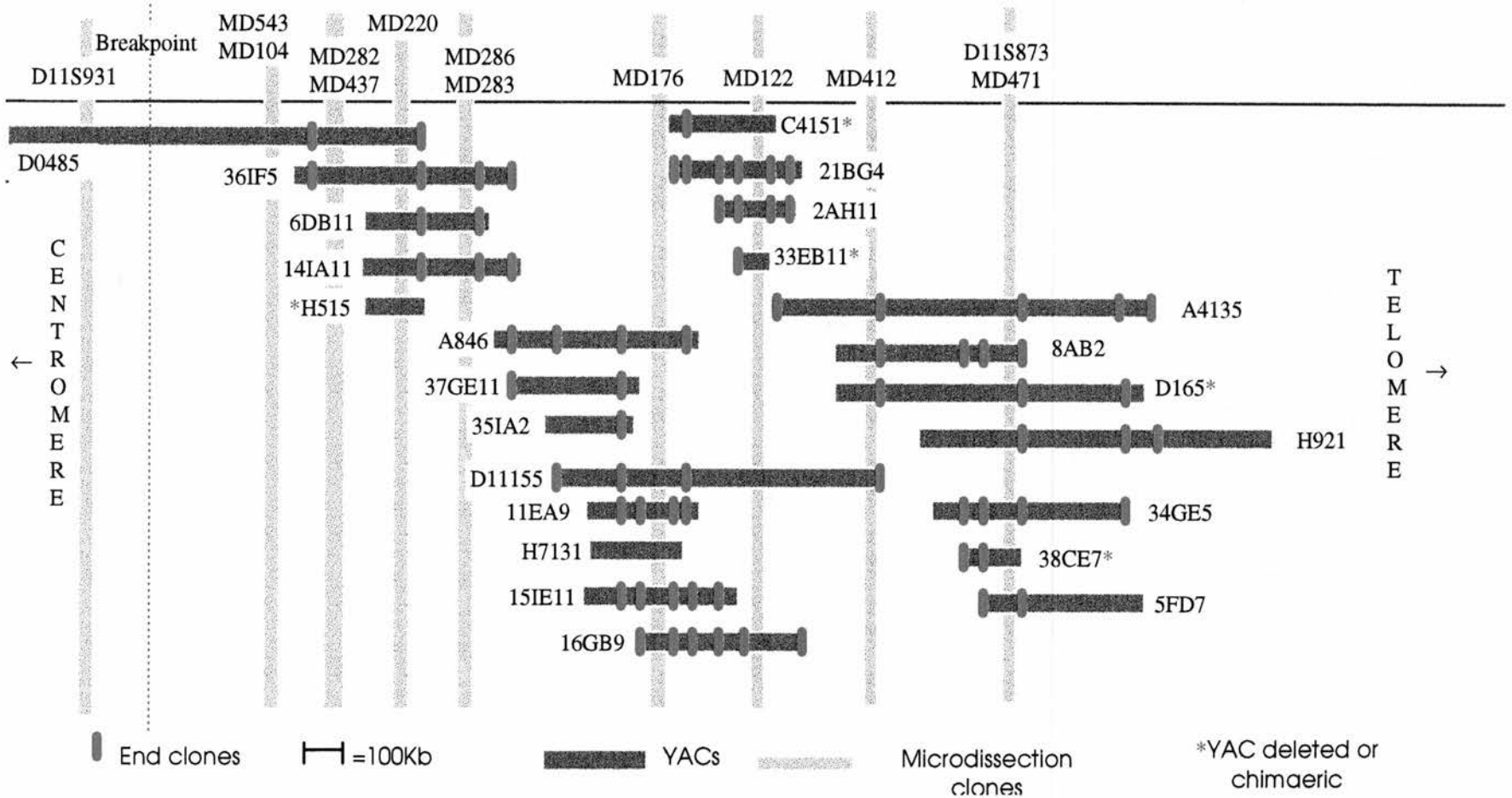


Figure 3.2: Chromosome 11 YAC contiguous clone map (~3Mb)

Table 3.1; Details of the somatic cell hybrid mapping panel.

Somatic Cell Hybrid	Chromosomes	Background
MIS7.4[1]	Derived 1 translocation chromosome	Murine X63
MIS39.8[1]	Derived 11 translocation chromosome	Murine X63
MAR1[1]	Derived 11 translocation chromosome	Murine RAG
MAR12[1]	Derived 1 translocation chromosome	Murine RAG
MAFLI[1]	B lymphoblastoid cell line from t(1:11)(q43,q21) translocation patient	Human
J1CL4[1]	Intact human chromosome 11 as sole human component	Chinese hamster ovary cell line WG3H
A9	Murine cell line	Murine
E67.1[2]	Single fragment of human chromosome 11	Murine C127
E67.4[2]	Single fragment of human chromosome 11	Murine C127
WJX3[1]	X-irradiation hybrids made from J1CL4	Chinese hamster ovary cell line WG3H
WJX7[1]	X-irradiation hybrids made from J1CL4	Chinese hamster ovary cell line WG3H
WJX7[1]	X-irradiation hybrids made from J1CL4	Chinese hamster ovary cell line WG3H
WJX11[1]	X-irradiation hybrids made from J1CL4	Chinese hamster ovary cell line WG3H
WG3H	Chinese hamster ovary cell line	Hamster
C127	Murine cell line	Murine
RAG	Murine cell line	Murine

[1]Fletcher et al 1993 [2]Porteous et al 1986

The somatic cell hybrid panels were made in a similar fashion to that of the YAC hybridisation panels, being cut with restriction enzymes *EcoR1* or *HindIII*, run on 1% agarose gels and Southern blotted (see section 2.8). On each somatic cell hybrid panel a total human genomic DNA control was included along with mouse only and hamster only cell lines in order that human specific bands could be identified.

3.2.2 Hybridisation Probe Production

The cDNA sequence for ACTN 2 and ACTN 3 has been published (Beggs et al 1992) and is available in the GENBANK database (HGMP Resource centre, Hinxton Hall, Cambridge: accession numbers: M86406 and M86407) However, little information was available on intron/exon boundaries in these two genes. Two intron/exon boundaries had been described in ACTN 2 as a result of the identification of a CA dinucleotide repeat polymorphism, described by Beggs et al (1992) and were available in the GENBANK database (accession number: M86804). No such information was available for ACTN 3.

Oligonucleotide primers were designed to the 3' untranslated regions (3' UTR) of the two α-actinin genes ACTN 2 and ACTN 3. 3' UTRs are usually more divergent between genes than coding regions and BLAST (Altschul 1990, Pearson and Lipman 1988) database searches indicated that these regions were less conserved in both the actinin genes which are 90% similar in cDNA sequence (Beggs et al 1992). By designing primers from the 3'UTR region, problems of cross hybridisation between the two actinin probes and between other members of the spectrin gene super family, which are highly homologous, were minimised.

PCR was done using 100ng of total human genomic DNA as template (see section 2.13 for description of PCR procedure). The oligonucleotide primers designed for ACTN 2 were; J101; 5'-GCTTCTGTAATCACTCATCCC-3' and J102; 5'-AGGATACTGGTTTCTGACTTG-3'. Those designed for ACTN 3 were J103; 5'-GAGACTGACACGACTGAGCAA-3' and J014; 5'-TGGCTGGCTTTTCTCTTAGGC-3'. Conditions used in the PCR reaction are detailed in Table 3.2. As the published 3'UTR sequence for ACTN 3 is only 133bp long, one of the oligonucleotide primers (J104) was designed within the 3'untranslated region and the other within the coding sequence (J103).

Table 3.2 PCR conditions for amplification of ACTN 2 and ACTN 3 probes (PCRs were carried out on Hybaid Omnigene PCR machines).

	Denature	Annealing	Extension
Cycle 1	94°C/2mins	55°C/30secs	72°C/1min
Cycle 2-29	94°C/30secs	55°C/30secs	72°C/1min
Cycle 30	94°C/30secs	55°C/30secs	72°C/2mins

The expected product sizes for ACTN 2 and ACTN 3 were 913bp and 316bp respectively. The ACTN 2 product was of the expected size but the ACTN 3 probe was slightly larger than expected at approximately 400bp (Figure 3.3). Both the PCR products were sequenced by single pass automated fluorescent sequencing (see section 2.14). The results showed the ACTN 2 probe to be as predicted from the published 3' untranslated region sequence. The ACTN 3 probe however was found to have an additional 80bp inserted into the probe sequence (Figure 3.4) at position 2565 of the published cDNA sequence. Analysis of this sequence showed that it is flanked by splice consensus sequences (Mount et al 1982) and therefore these primers span a small intron in genomic DNA.

Comparison of the ACTN 2 and 3 probe sequences showed that the two probes are not homologous and should not cross react (data not shown).

Several PCR products for ACTN 2 and ACTN 3 were pooled together and run out on a low melting point agarose gel. Bands of the appropriate size were then cut out and the DNA purified using a Geneclean II kit (see section 2.7.5). This DNA was used as a probe to hybridise to the YAC and somatic cell hybrid panels.

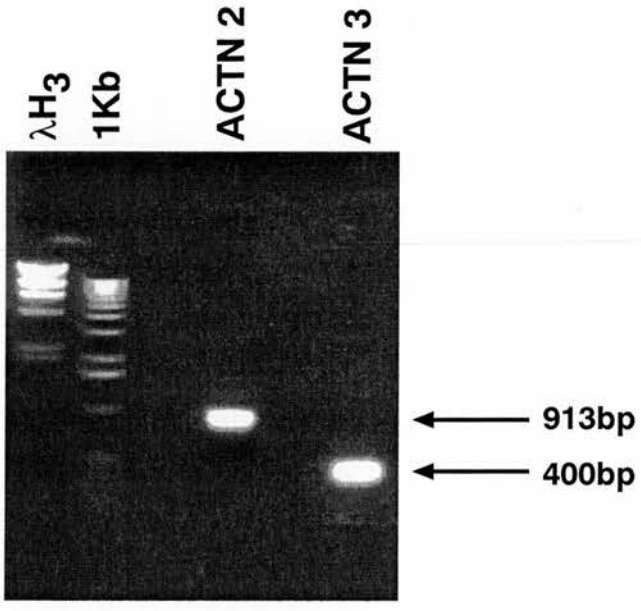


Figure 3.3: ACTN 2 and ACTN 3 PCR probes from 3'untranslated region of the genes. Expected product sizes were ACTN 2 - 913bp and ACTN 3 - 316bp.

Figure 3.4a: Location of ACTN 3 intron within the cDNA sequence. The position of the start of the 3'UTR is marked with an arrow. Primers used to amplify the probe from genomic DNA are identified (J103 and J104). The position of the intron is marked in red.

```
2451  CGCAGCTGGG GTGGTGACCT TCCAGGCCTT CATAGACTTC ATGACCCGAG
                      J103
2501  AGACAGCCGA GACTGACACG ACTGAGCAAG TTGTAGCTTC CTTCAAGATC

2551  TTGGCAGGAG ACAAGAACTA CATCACCCCC GAGGAGCTGC GGCGCGAGCT
                      INTRON
2601  CCCTGCCAAG CAGGCCGAGT ACTGCATCCG CCGTATGGTG CCCTACAAGG

2651  CATCCCCGGC CCCGGCTGGA GCCCTGGACT ACGTGGCCTT CTCCAGTGCC
                                    3'UTR
2701  CTCTATGGGG AGAGCGACCT TTGACCCCAA CCACTGAGGT TCTCTATGCA

2751  AGATGGAGAG AGGATGCACC CTGTGGCTGA TCCCATCCGT CCCTCGGAGC

2801  AAGGGCCTAA GAGAAAAGCC AGCCAAGTGC TTCTGAATAA AGATCCCTCT
                J104
2851  CTGGGTCA
```

Figure 3.4b: Sequence of the ACTN 3 PCR probe, indicating the intron sequence. The intron sequence is underlined with dotted lines indicating the splice consensus sequences.

```
  1  CTyCCTTCAA GATCTTGGCA GGAGACAAGG TGAGTCCCAA GCGGTGGAGG

 51  GGCTGGTGGG CTAGGGCAGG GCACGGGACC TGTGGGTCCT CAACGCCTCT

101  TCTCCCCAGA ACTACATCAC CCCCGAGGAG CTGCGGCGCG AGCTCCCTGC

151  CAAGCAGGCC GAGTACTGCA TCCGCCGTAT GGTGCCCTAC AAGGGATCCG

201  GGGCCCCGGC TGGAGCCCTG GACTACGTGG CCTTCTCCAG TGCCCTCTAT

251  GGGGAGAGCG ACCTTTGACC CCAACCACTG AGGTTCTCTA TGCAAGATGG

301  AGAGAAGATG CACCCTGTGG CTGATCCCAT CCGTCCCTCg GAGCAAGGCA
```

3.2.3 Southern Blot Hybridisations to Chromosome 1 and 11 YAC and Somatic Cell Hybrid Panels

The ACTN 2 3'UTR probe was hybridised to a filter containing YAC DNA from the chromosome 1 contig and the somatic cell hybrids all of which had been cut with *HindIII* restriction enzyme (see section 3.2.1). The results from this hybridisation (Figure 3.5) showed a band corresponding to approximately 1Kb in size in the A9 (human chromosome 1 only), MIS39.8 and MAR1(derived chromosome 11containing hybrids) somatic cell hybrids and in total human genomic DNA (THG). This band does appear to be slightly larger in size in the two hybrids MIS39 and MAR1 than in A9 and total human genomic DNA. This is most likely due to the greater amount of DNA loaded on the Southern blot for MIS39.8 and MAR1 in comparison to A9 and total human DNA lanes (see gel picture Figure 3.5). The large amount of DNA is likely to cause slight retardation of the DNA as it runs through the gel resulting in the slight band size difference. There is also a difference in the intensity of the bands seen with the MIS39.8 and MAR1 hybrids being less strong than expected given the difference in the amount of DNA loaded on the blot. This is a reflection of the low copy number in the hybrids in comparison to total human genomic DNA. A slightly smaller band at approximately 500bp was seen in total human genomic DNA (THG) and in MAR 1 hybrid but is only very faintly visible in MIS39.8 and A9 hybrids. No bands were seen in the derived 1 hybrids MIS7.4 and MAR12 or in the chromosome 11only hybrid J1CL4 (indicating that this probe does not cross hybridise with ACTN 3 on chromosome 11)(See Figure 3.5). Similarly, no bands were seen in the hamster or mouse only cell lines, indicating that this probe does not hybridise across these species. To confirm that the ACTN 2 probe is contained in these hybrids PCR analysis was done using 200ng of DNA in the same PCR reaction that was used to amplify the ACTN 2 probe (see Table 3.2). The PCR results are displayed in Figure 3.6. The results confirm those found in the Southern blot analysis.

None of the YACs contained a band of 1Kb in size (refer to right hand side of figure 3.5) but YAC 36DH11 did contain a band which was slightly smaller than this, at approximately 800bp. This band was not seen in any of the other YACs in the contig which are known to overlap with 36DH11(31AH8, 757F9, 851G8) indicating that, if this band is real, this YAC may be chimaeric. A band of the same size to that seen in YAC 36DH11was also present in the two lanes which contained size marker DNA indicating that this band may be a result of plasmid contamination in the positive

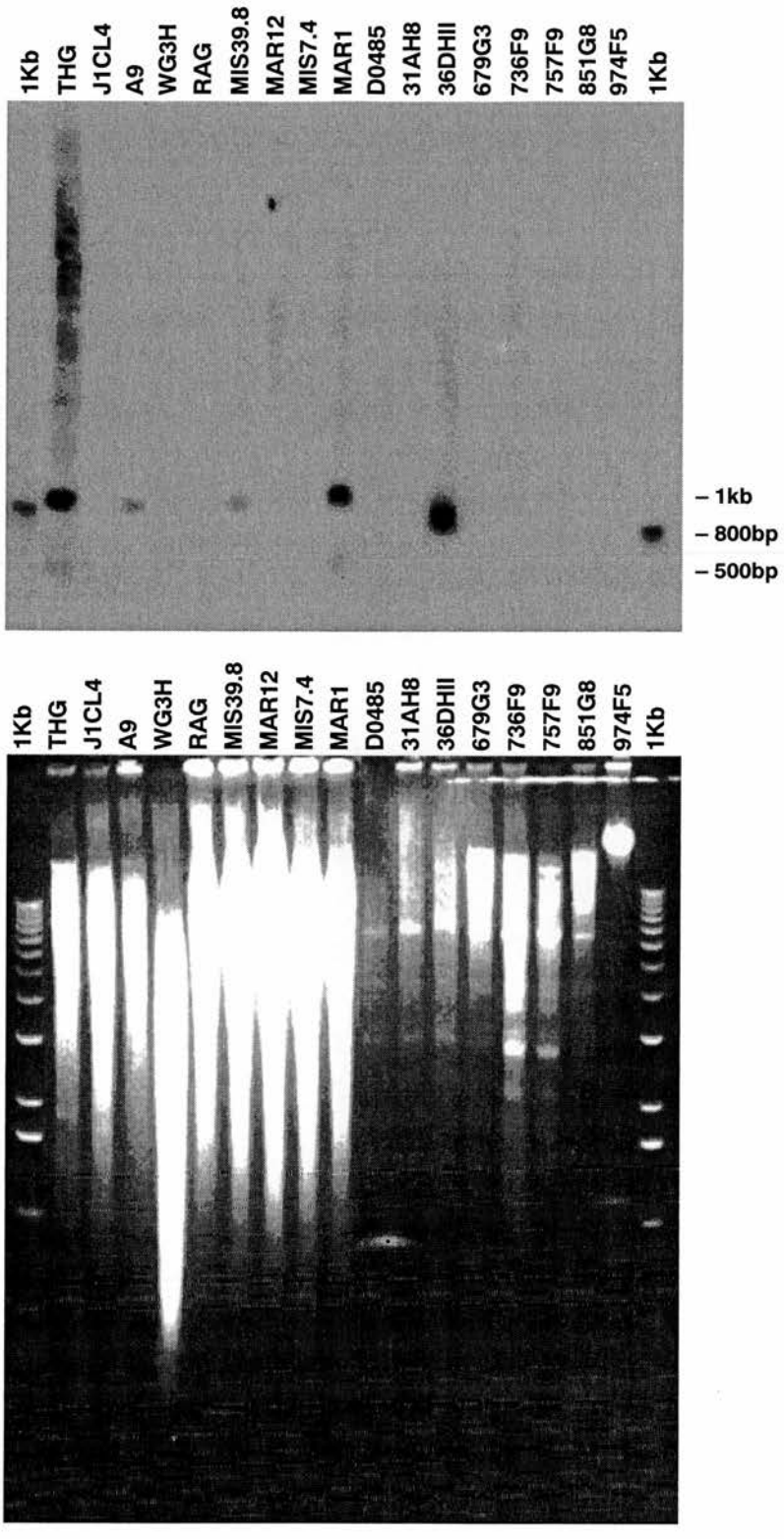


Figure 3.5: Southern blot hybridisation analysis of ACTN 2 3'UTR probe hybridised to the chromosome 1 YAC and somatic cell hybrid panel cut with restriction enzyme *HindIII* (THG = total human genomic DNA)

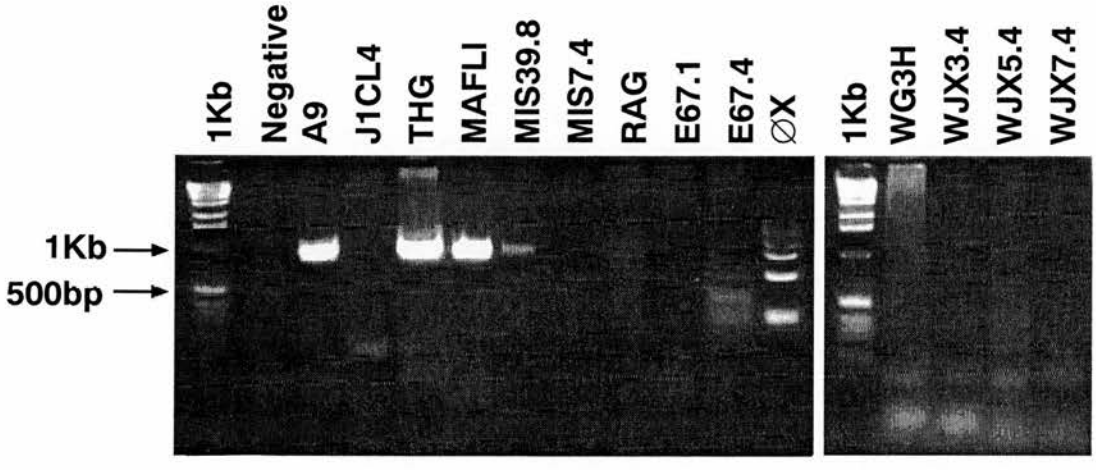


Figure 3.6: PCR analysis showing amplification of ACTN 2 3'UTR from the somatic cell hybrids.
THG=total human genomic DNA.

YAC since size markers contain plasmid DNA. To determine if this band in YAC 36DH11 was real or not PCR was done on the YAC DNA to try to amplify a band corresponding to the ACTN 2 probe. PCR conditions were the same as for ACTN 2 probe production and 200ng of YAC DNA was used in the PCR reaction. Total human genomic DNA and the human chromosome 1 only hybrid, A9, were included as positive control samples. Results from the PCR (Figure 3.7) showed that 36DH11 did not contain a band corresponding to the ACTN 2 gene. A band was seen in the positive control samples A9 and total genomic human DNA but in none of the chromosome 1 YACs (Figure 3.7). It is probable that the band seen on hybridisation in YAC 36DH11 is the result of plasmid contamination of this YAC DNA and does not represent the ACTN 2 gene. It is likely therefore that the ACTN 2 gene is not contained in any of the chromosome 1 YACs. Since the chromosome 1 YAC contig spans approximately 2Mb around the breakpoint the ACTN 2 gene most probably lies out-with this region.

The ACTN 3 3'UTR probe was hybridised to chromosome 11 YAC filters which had been cut with restriction enzyme *EcoR1*. The results (not shown) demonstrated that the ACTN 3 probe was not contained in any of the YACs in the chromosome 11contig. Only 1 YAC, D0485, crosses the translocation breakpoint and therefore extends proximal to the breakpoint, all the other YACs in the contig are distal to the breakpoint. Since YAC D0485 only extends 500Kb proximal to the breakpoint it is possible that the ACTN 3 gene could lie just beyond this YAC, still being within a reasonable distance of the breakpoint. To exclude this possibility high resolution mapping employing somatic cell hybrids was carried out.

The ACTN 3 probe Southern blot analysis to an *EcoR1* somatic cell hybrid panel showed that several *EcoR1* fragments hybridised with the ACTN 3 probe in total human genomic DNA and that hamster and mouse fragments also hybridised with this probe (Figure 3.8). The total human genomic DNA appears to contain three major bands on a background smear. The smear is likely to be due to the greater complexity of the genomic DNA and also possibly due to repetitive elements. The largest of these bands (approximately 23Kb) also appears in the chromosome 11 only hybrid J1CL4 and the derived 11 hybrid MIS39.8 as well as in E67.1 and E67.4 which contain fragments of chromosome 11. Two bands (~10Kb and ~4Kb) were also seen in all the hybrids which were in a mouse background and the mouse only

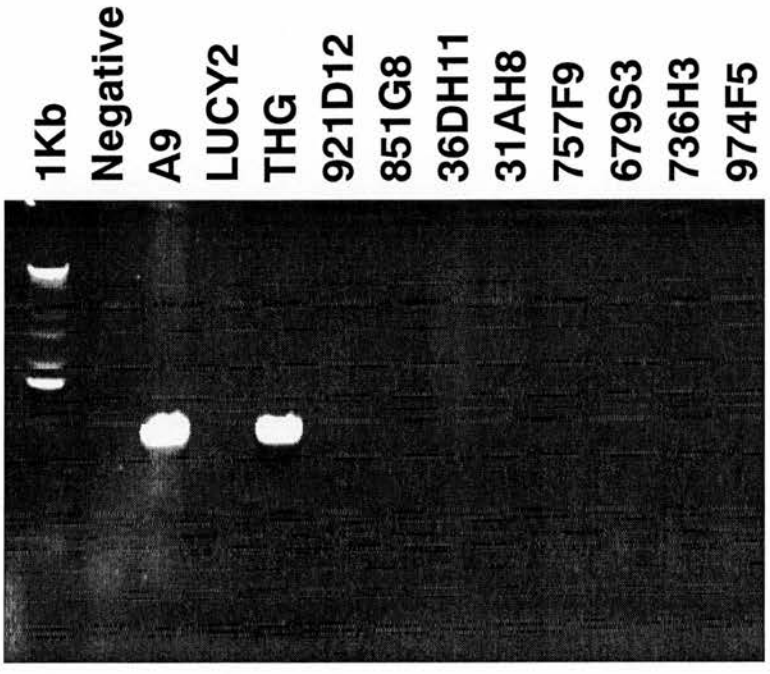


Figure 3.7: PCR analysis showing amplification of ACTN 2 3'UTR from the chromosome 1 YACs

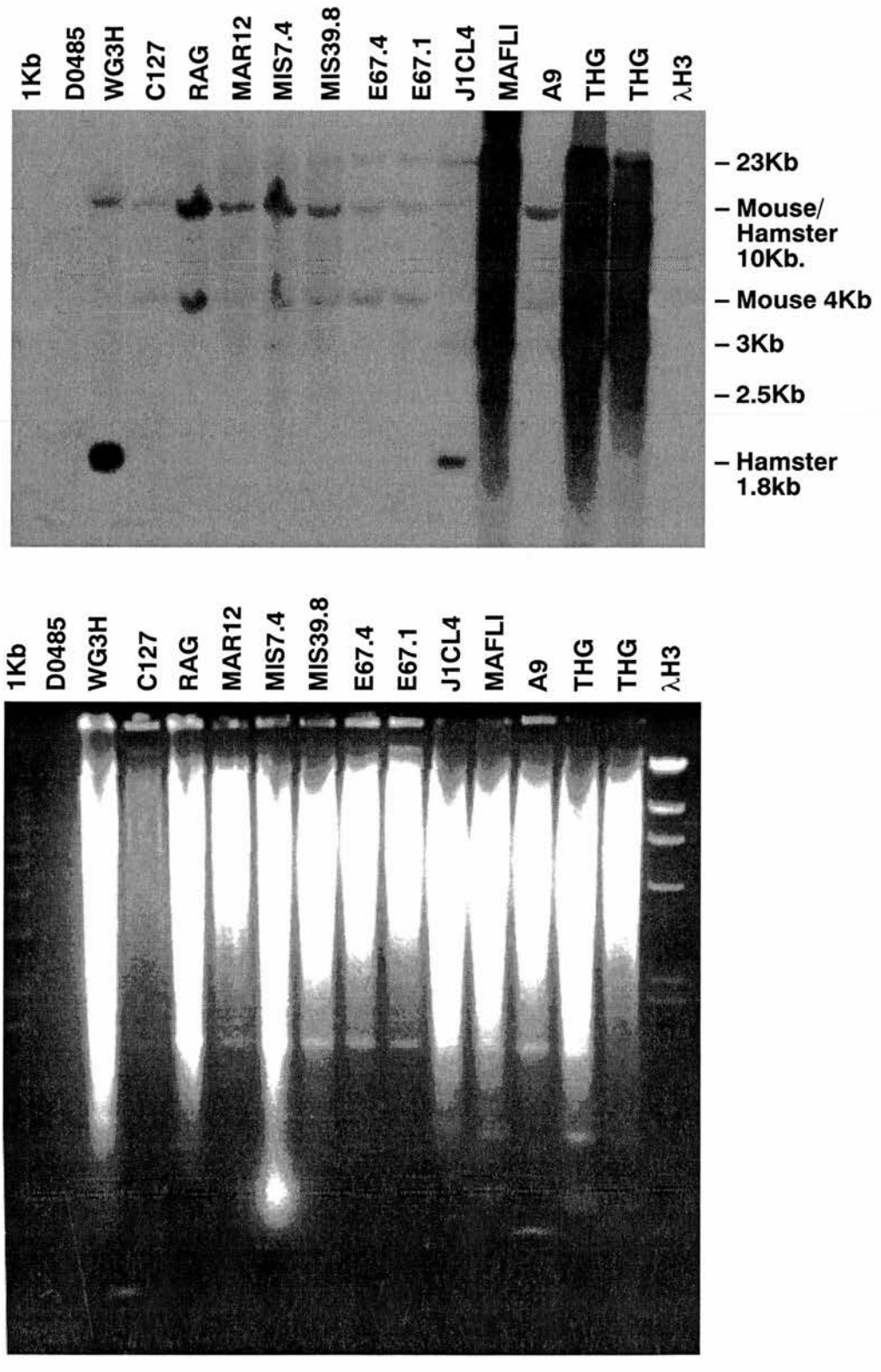


Figure 3.8: Southern blot hybridisation analysis of ACTN 3 3'UTR probe hybridised to the chromosome 11YAC and the somatic cell hybrid panel cut with restriction enzyme *EcoRI.* (THG = total human genomic DNA).

hybrid C127, indicating that the ACTN 3 probe cross hybridises to homologous genes in the mouse. Two bands of slightly different size to those seen in the mouse were also seen in J1CL4 and the hamster background hybrid WG3H (~10Kb and ~1.8Kb). The only band which was seen in the chromosome 11 hybrids and total human genomic DNA which didn't correspond to either a mouse or hamster band was the ~23Kb band and was therefore deemed the human ACTN 3 band. A slightly smaller band to that designated as the human ACTN 3 band is seen in the hybrid MAR12. This hybrid contains other chromosomes as well as the derived 1 translocation chromosome (although it does not contain normal chromosomes 1 or 11). It is therefore possible that this band represents a related family member which resides on a different chromosome.

A PCR assay using primers J103 and J104 was done in order to confirm the Southern blot analysis (Figure 3.9). 200ng of DNA was used in the reaction which was carried out as described in Table 3.2. Strong bands of the correct size were seen in J1CL4, total human genomic DNA, MIS39.8, E67.1, E67.4, WJX3.4 and MAFLI (translocation patient total genomic DNA). No other bands were seen in any of the other hybrids confirming the results obtained from the Southern blot analysis.

Since the somatic cell hybrids have already been characterised in terms of chromosome 11 markers (Evans et al 1995) it was possible to map the ACTN 3 gene to a defined interval on chromosome 11 (See Table 3.3). This interval was at the GST PI Locus corresponding to a map position at 11q13. The GST PI locus is the only locus in which E67.1, E67.4, MIS39.8 and WJX 3.4 are all positive and the other hybrids being negative.

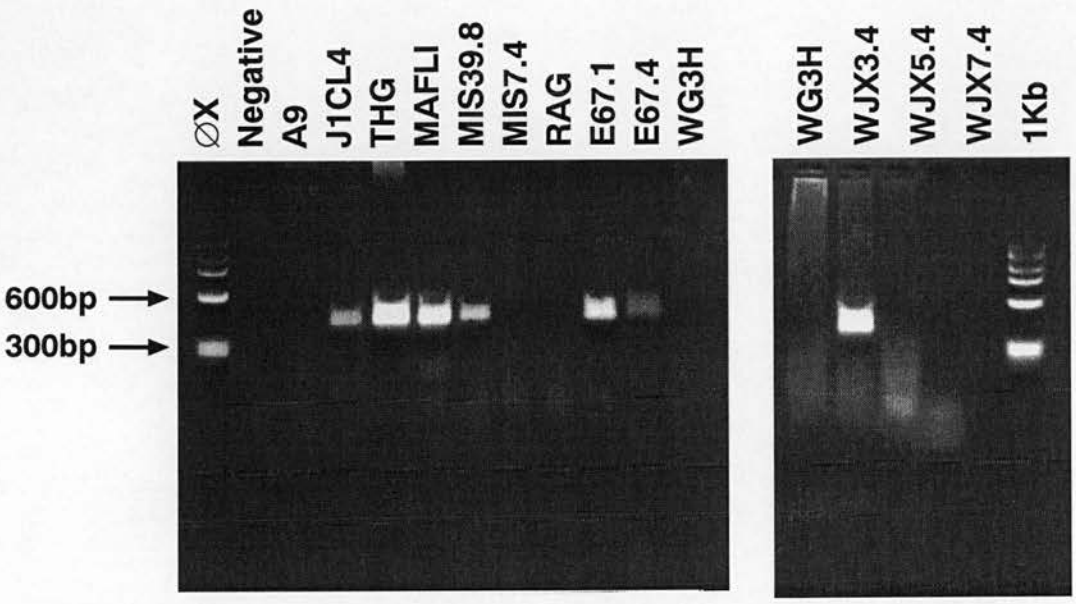


Figure 3.9: PCR analysis showing amplification of ACTN 3 3'UTR from the somatic cell hybrids. (THG = total human genomic DNA).

Table 3.3: Chromosome 11 marker analysis of somatic cell hybrids (reproduced from Evans 1994). + indicates the presence of a marker in the cell line, - indicates its absence and ND indicates not determined. The GST PI locus where the ACTN 3 gene is predicted to map from the Southern Blot analysis is in bold typeface and underlined.

	SOMATIC CELL HYBRIDS								
LOCUS	**MIS 7.4**	**MIS 39.8**	**CF52**	**WJX 3.4**	**WJX 5.4**	**WJX 7.4**	**WJX 11.2**	**E67.1**	**E67.4**
HBB	-	+	ND	-	+	-	-	+	-
D11S87	-	+	ND	-	-	-	-	ND	ND
TYR LI	-	+	-	+	+	-	-	-	+
PGA	-	+	-	+	-	+	-	+	-
ROM1	-	+	-	+	-	+	-	-	-
MDU 1	-	+	-	+	-	+	-	-	-
GST PI	**-**	**+**	**-**	**+**	**-**	**-**	**-**	**+**	**+**
ADROA2	-	+	ND	+	-	-	-	+	-
FGF 4	-	+	-	+	-	-	-	-	-
D11S527	-	+	+	+	-	-	-	-	-
D11S533	-	+	+	+	-	+	-	-	+
OMP	-	+	+	+	-	+	-	-	+
TYR	-	+	+	-	-	+	+	-	+
CLG	+	-	+	-	+	-	+	-	+
STMY1	+	-	+	-	+	-	+	-	+
D11S385	+	-	+	-	+	-	+	-	+
NCAM	+	-	+	-	-	+	+	-	+
DRD2	+	-	+	-	-	+	+	-	-
D11S351	+	-	+	-	+	-	+	-	-
THY1	+	-	+	-	+	-	+	ND	ND

3.3 Discussion

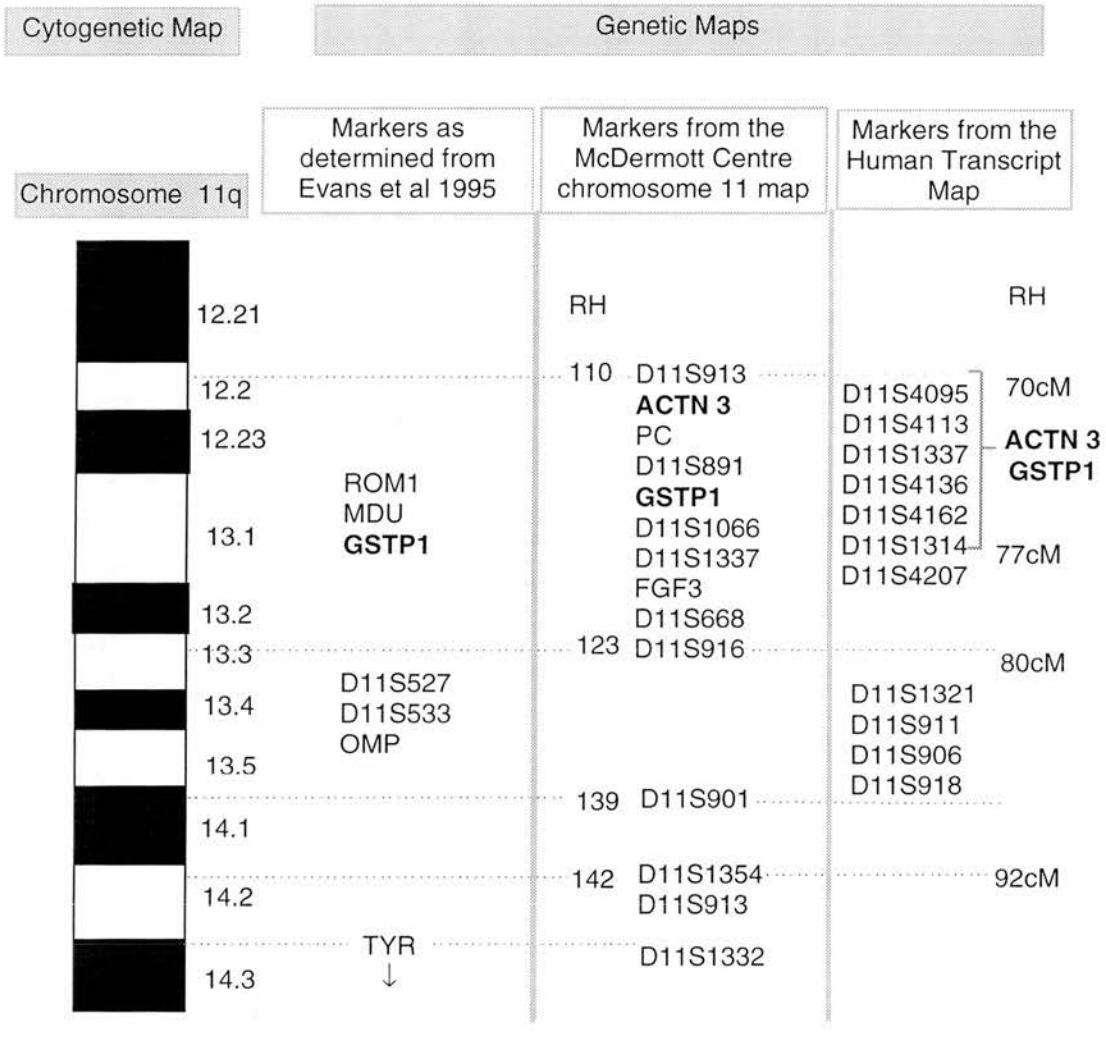
The Southern blot hybridisations and PCR analysis have shown that both ACTN 2 and ACTN 3 lie out-with the region of the YAC contigs spanning the translocation breakpoints on chromosome 1 and 11. The chromosome 1 contig spans approximately 1Mb on either side of the breakpoint and since ACTN 2 has been shown to most probably lie out-with these YACs, it can be deemed to be greater than 1Mb away from the translocation breakpoint on chromosome 1. Since there are several other identified cDNAs which are closer to the breakpoint, ACTN 2 was considered to be too far from the breakpoint to be further analysed as a candidate gene involved in the psychiatric diagnosis in the K26 t(1:11) translocation family. Southern blot hybridisation of ACTN 2 with somatic cell hybrids confirmed that ACTN 2 map position is distal to the translocation breakpoint as it was located on the derived 11 chromosome. It is not possible to refine this mapping position further as characterised chromosome 1 somatic cell hybrids are not available to us. ACTN 2 has however been mapped to the interval D1S439-D1S446 (246-256cM on radiation hybrid map) in the Human Transcript Map (Schuler et al 1996, http://www.ncbi.nlm.nih.gov/SCIENCE96/). The chromosome 1 YAC contig is contained within this region in the interval D1S225-D1S459 (249-251cM - the breakpoint being approximately half way between these markers).
Since we know that ACTN 2 is not contained within the YAC contig (i.e. between markers D1S225-D1S459 - 249-251cM) and that it maps proximal to the breakpoint, it probably lies in the region between markers D1S459 and D1S446 (251-256cM).

Southern blot hybridisation analysis and PCR similarly showed that ACTN 3 did not map to the chromosome 11 YAC contig. This result would place this gene greater than 2Mb distal to the breakpoint but since only 1 YAC (D0485) spans the chromosome 11 breakpoint this gene could lie just beyond this, at a distance of greater than 500Kb or so proximal to the breakpoint. Southern blot analysis using somatic cell hybrids placed ACTN 3 in the region of the GST PI locus at 11q13, some considerable distance proximal to the chromosome 11 breakpoint.
Quackenbush et al (1995) produced an STS content map of human chromosome 11 in which they localised 910 chromosome 11 specific YAC clones. They mapped ACTN 3 to chromosome 11q12-13.1 by STS and radiation hybrid mapping, placing it between markers D11S913 and D11S916 at position 113 on their radiation hybrid

map. The translocation breakpoint is flanked by marker TYR on chromosome 11 which is also on the radiation hybrid map produced by Quackenbush residing at position 11q14.2 (position 146 on the radiation hybrid map). From their STS YAC contig map they estimate the distance between these two markers to be greater than 15Mb, some considerable distance from the breakpoint (although it must be emphasised that this distance is by no means accurate but serves to demonstrate the large intervening distance between the ACTN 3 gene and the breakpoint).
The GST PI marker is also positioned on their radiation hybrid map at approximately position 113 and therefore lies extremely close to where they have positioned ACTN 3. The Human Transcript map (Schuler et al 1996) has also positioned both GSTP1 and ACTN 3 on their radiation hybrid map at between markers D11S913 and D11S1314 (70-77cM)(see Figure 3.10). These results therefore confirm the validity of the somatic cell hybrid mapping approach used to map ACTN 3.

The two α-actinin genes ACTN 2 and ACTN 3 were considered as candidate genes for involvement in the psychiatric diagnosis in a family in which schizophrenia and associated disorders cosegregate with a t(1:11) balanced translocation on the basis of location in close proximity to the breakpoint and biological function. As these two homologous genes lie in close proximity to the translocation breakpoint they may also provide a mechanistic explanation of the translocation event by illegitimate cross over.
High resolution mapping of these two genes using YAC and somatic cell hybrid panels from the region indicated that these two genes map some considerable distance from the breakpoint making it highly improbable that they are involved in the psychiatric diagnosis in this family.
Exclusion of ACTN 2 and 3 from the search for genes involved in the psychiatric diagnosis in the translocation family will allow more effort to be directed to candidate genes which map in close vicinity to the breakpoint.

Figure 3.10 Cytogenetic, genetic and radiation hybrid maps from Evans et al (1995), the McDermott Centre for human growth and development (http://mcdermott.swmed.edu/) and the Human Transcript Map (http://www.ncbi.nlm.nih.gov/SCIENCE96/).



RH - Radiation Hybrid Map

Chapter Four

High Resolution Mapping of Expressed Sequence Tagged Sites on Chromosome 11

Introduction

In 1995 Rosier et al regionally assigned 68 new human gene transcripts (ESTs) to 25 sub-regions of chromosome 11. These transcripts were originally obtained from a collection of partial cDNA sequences derived from skeletal muscle and infant brain cDNA libraries established at Généthon by the Genexpress program. These ESTs had been previously assigned to chromosome 11 using a panel of non-chromosomic somatic cell hybrids. Using a collection of somatic cell hybrids containing different deletion fragments of chromosome 11 (which had previously been characterised by cytogenetic data, reference markers and by their microsatellite content), assignment of the ESTs to the 25 subintervals of chromosome 11 was achieved (Rosier et al 1995). Among the somatic cell hybrids used to regionally assign these EST were the hybrids MAR1and MAR12, containing the derived 11 and derived 1 translocation chromosomes respectively, from the K26 pedigree.

Rosier et al (1995) identified 69 diseases which were genetically linked to chromosome 11. For 31 of these diseases no responsible gene had been identified resulting in them being termed "orphan pathologies". The regions of chromosome 11 to which these orphan pathologies had been linked were used to generate a disease map of chromosome 11 which was integrated with the cytogenetic, genetic and genic maps. Each of the 25 subintervals defined on chromosome 11 has at least one EST contained within it so that each orphan pathology has at least one EST localised to the region close to it, thus providing positional candidate gene(s) for that disease which could then be investigated. The region of chromosome 11 in which our translocation, associated with schizophrenia, resides is indicated on this integrated map and falls into subinterval 21-22 of the 25 subintervals as defined by their somatic cell hybrids. Region 21 lies above the translocation breakpoint as determined by the translocation hybrid MAR1 which is represented on the hybrid panel and extends from11q14.3-q21(D11S1789-D11S1780). This region contains 4 ESTs. Region 22 lies below the translocation breakpoint as defined by MAR12, extending from 11q21-23.2 (D11S1342-D11S1327) and contains 16 ESTs (see Figure 4.1).

The position of the translocation breakpoint on chromosome 11 in the Rosier et al paper, as indicated by the positions of MAR1and MAR12 hybrids, is 11q21. The position of the breakpoint has been further refined since the publication of this paper and actually resides at 11q14.3.

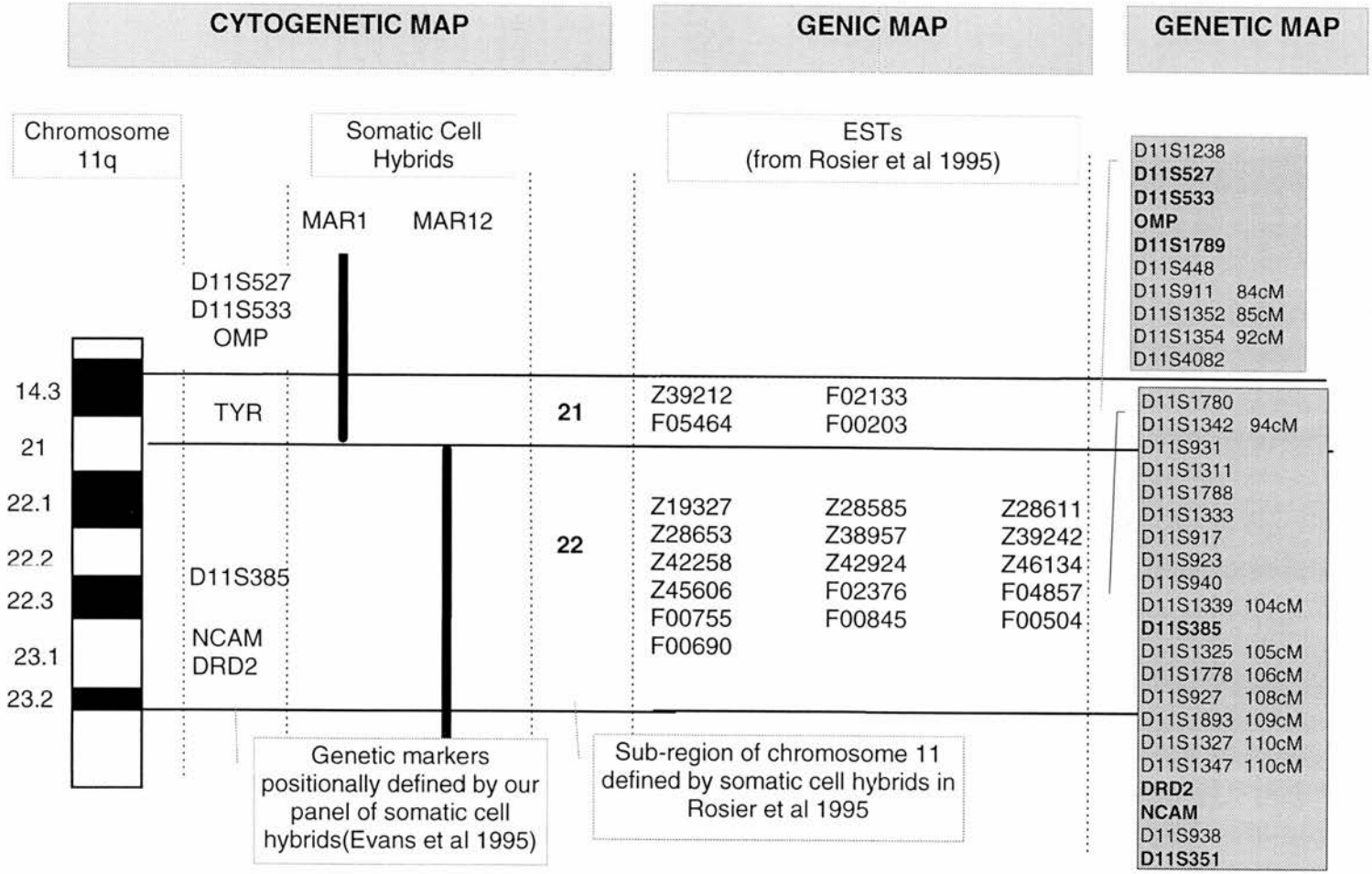


Figure 4.1: Integrated view of cytogenetic, genic and genetic map of a section of chromosome 11q. Data combined from mapping data from Evans et al (1995), Rosier et al (1995), Human Transcript Map, and McDermott Centre for Human Growth and Development chromosome 11 integrated map (http//mcdermott.swmed.edu/). Centimorgan (cM) values from Human Transcript map radiation hybrid map.

These newly identified transcripts which flank the chromosome 11 breakpoint are potential candidate genes for involvement in the psychiatric diagnosis in the K26 pedigree. In order to further determine the candidacy of these gene transcripts it was necessary to determine their proximity to the breakpoint and to establish if any of them crossed the translocation breakpoint. To this end, high resolution mapping of these ESTs, using chromosome 11 YACs and somatic cell hybrids was undertaken.

4.2 Mapping of ESTs to chromosome 11 YACs:

Of the 20 ESTs (See Table 4.1) which were mapped in the region around the translocation breakpoint by Rosier et al (1995) three ESTs were identified by sequence database searches as corresponding to known genes, namely mitochondrial acetoacetyl-CoA thiolase (F00504), small nuclear ribonucleotide polypeptide C (F05464) and alpha B-crystallin (F00690). The other 17 ESTs were of unknown sequence identity in the database (Rosier et al 1995). The alpha B-crystallin gene (F00690) had previously been mapped to 11q22.3-q23.1 and the mitochondrial acetoacetyl-CoA thiolase gene (F00504) had also been previously mapped to 11q22.3-q23.1. These two EST were therefore not looked at further due to their remote location from the breakpoint.

The 18 ESTs which showed no sequence identity in the database and EST F05464 were mapped to the chromosome 11 YAC contig panel. Since no primer pairs were described in the paper for amplification of the ESTs, the corresponding cDNA clones (which were listed in the paper) were ordered from the IMAGE consortium (Lennon et al 1996) clones held at HGMP, Hinxton Hall, Cambridge. Only 10 clones were available (see Table 4.1) from this source and the cDNA clone insert was used as a probe in Southern blot hybridisation analysis on the chromosome 11 YACs. The remaining ESTs for which no cDNA clones were available were mapped to the chromosome 11 YACs by PCR analysis using unique primers which were designed to the published EST sequence.

Table 4.1: Chromosome 11 ESTs residing in subintervals 21 and 22 as determined from somatic cell hybrid mapping by Rosier et al 1995.

ESTs	Chromosome 11 Sub-interval (Rosier et al)	cDNA Clone Available	Mapping Method	Database Identity
F05464	21	YES	Hybridisation	small nuclear ribonucleotide polypeptide C
F02133	21	YES	Hybridisation	None
F00203	21	NO	PCR	None
Z39212	21	YES	Hybridisation	None
F00504	22	N/D	N/D	mitochondrial acetoacetyl-CoA thiolase
F00690	22	N/D	N/D	Alpha B-crystallin
F04857	22	YES	Hybridisation	None
F02376	22	YES	Hybridisation	None
Z38957	22	YES	Hybridisation	None
Z39242	22	YES	Hybridisation	None
Z42258	22	YES	Hybridisation and PCR	None
Z42924	22	YES	Hybridisation and PCR	None
Z46134	22	YES	Hybridisation and PCR	None
Z28525	22	NO	PCR	None
Z28611	22	NO	PCR	None
Z28653	22	NO	PCR	None
Z19327	22	NO	PCR	None
F00755	22	NO	PCR	None
F00845	22	NO	PCR	None
Z45606	22	NO	PCR	None

Sub-interval 21 corresponds to regional classification in the Rosier et al (1995) paper and is the region just proximal the chromosome 11 breakpoint. Sub-interval 22 corresponds to the region distal the translocation breakpoint on chromosome 11.

4.2.1 High resolution mapping of chromosome 11 ESTs by Southern Blot hybridisation Analysis

cDNA clones from which the ESTs were derived were obtained for EST's F05464, F02133, Z38957, F04857, F02376, Z39212, Z39242, Z42258, Z42924 and Z46134 from HGMP Resource Centre, Hinxton, Cambridge.

Probes were made from the cDNA clone insert by boiled colony PCRs (see section 2.13.2), employing primers designed to the Lafmid BA vector sequence (291; 5'-CAGGAAACAGCTATGAC-3'and 292; 5'-GTAAAACGACGGCCAGT-3'), into which the cDNA had been cloned. PCR products were run out on a 1% low melting point agarose gel. The product was cut out and the DNA purified using Geneclean kit (section 2.7.5). These probes were then used for hybridisation to the chromosome 11 YAC panels, which are described in section 3.2.1. All YAC panels contained chromosome 11 only hybrid J1Cl4 and total human genomic DNA as positive controls.

The results of these hybridisations indicated that 9 out of the 10 cDNA clones from which the ESTs were derived, were not contained within any of the YACs in chromosome 11 YAC contig. Negative results were reported only when positive bands were seen in total human genomic DNA and J1CL4 with all the YACs being negative. An example of such a result is shown in Figure 4.2. In the majority of the cDNA hybridisations a constant band, which differs in size from that seen in the positive controls, is seen in all of the YACs (see Figure 4.2). This band is approximately 8Kb (*EcoR1*) in size and appears on several of the YAC hybridisation blots with different probes. This band is not consistent with the vector bands of the Lafmid BA vector or with Yeast specific bands (as indicated by the inclusion of Lucy2 yeast only cell line in some of the panels). These bands were initially thought to have resulted from some contamination from the boiled colony at the PCR stage. However, they also appear when a PCR product from the 3' end of EST Z39242 prepared from Qiagen minipreped DNA was used as a probe (see Figure 4.3) and therefore may be the result of contamination in the YAC DNA perhaps resulting from the preparation of the DNA as it does not appear in the human hybrids or human genomic DNA. Alternatively, these bands could be the result of YAC vector contamination. These results indicate that these ESTs are not in close proximity to the translocation breakpoint since they do not reside within the YAC contig. This would place them at a distance of greater than 2Mb distal, and 800Kb proximal to the chromosome 11 breakpoint.

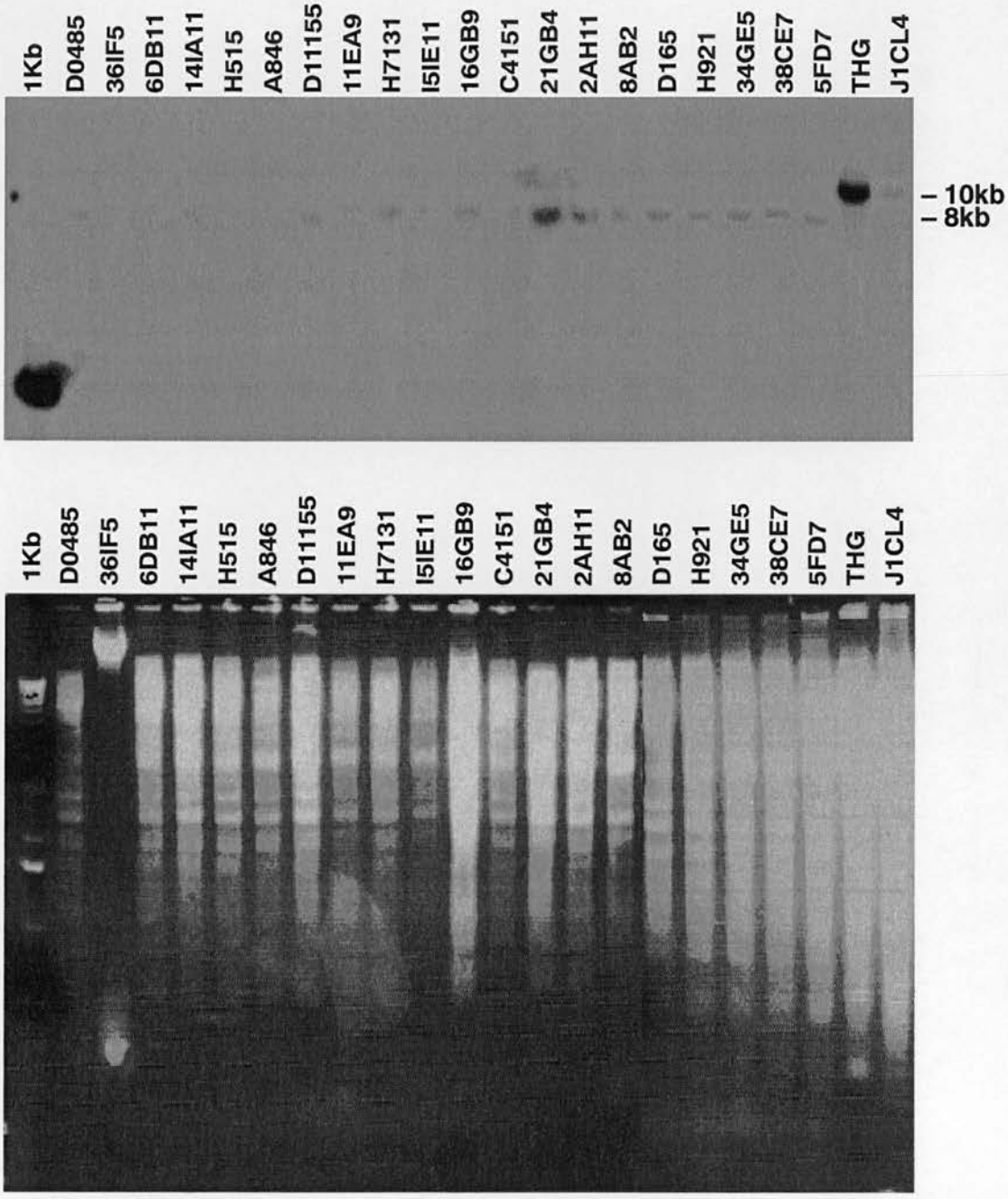


Figure 4.2: An example of Southern blot hybridisation analysis of the cDNA clone corresponding to EST F02133 onto chromosome 11 YAC panel cut with restriction enzyme *EcoRI*. Positive controls THG(total human genomic DNA) and J1 (chromosome 11 only hybrid) are included in the panel. A constant band at 8Kb is visible in all the YACs but the positive controls have a larger band only at approximately 10Kb.

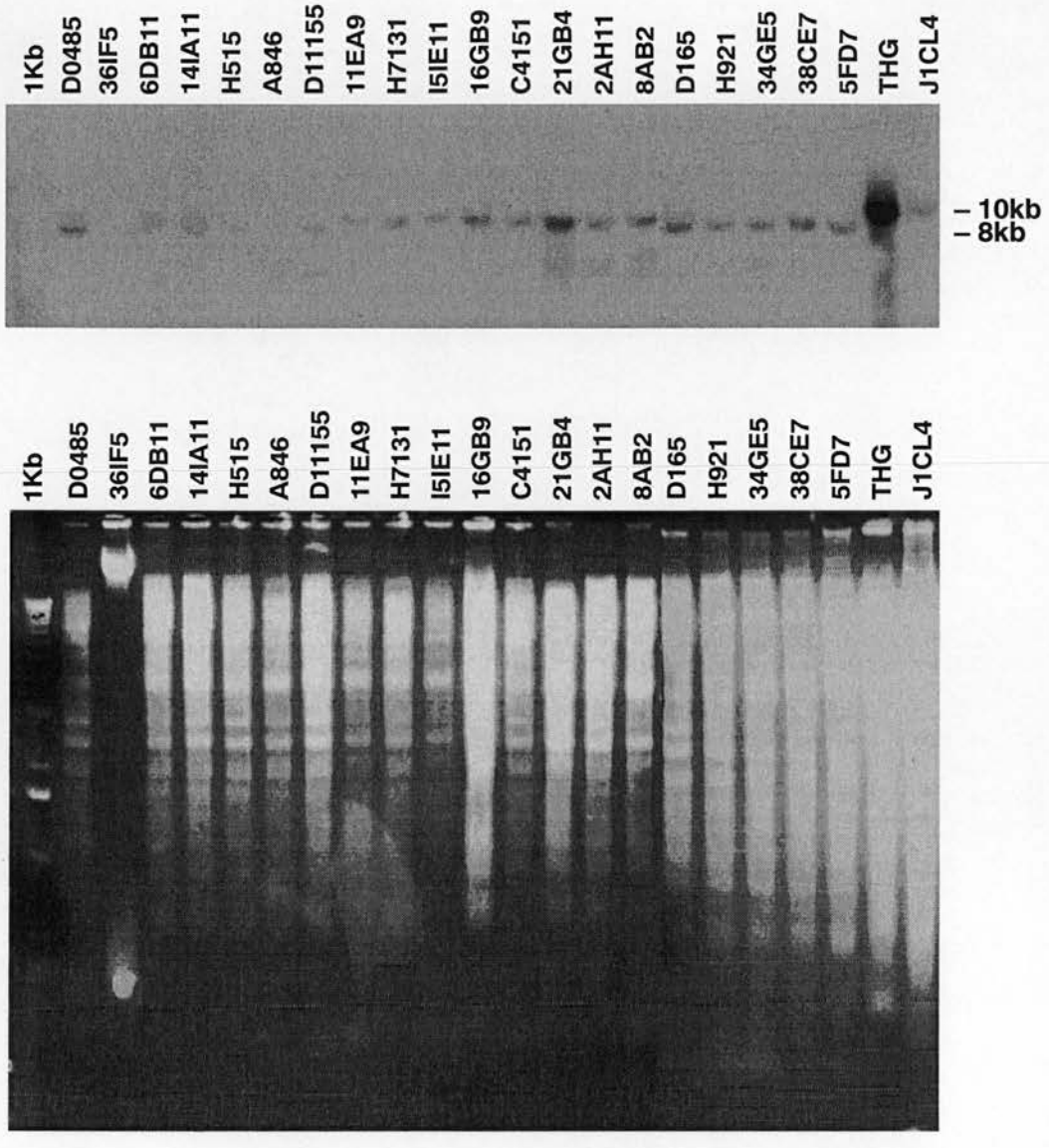


Figure 4.3: Southern blot hybridisation analysis of PCR product from 3' end of the cDNA clone corresponding to the EST Z39242 hybridised onto a chromosome 11 YAC panel cut with *EcoRI.*

The cDNA clone, from which the EST Z39242 was derived, appeared to contain bands in several of the YACs in the chromosome 11 contig, as well as in the positive controls J1CL4 and total human genomic DNA (Figure 4.4A and 4.4B). This clone was investigated in more detail.

cDNA clone corresponding to EST Z39242:

The cDNA clone from which the EST Z39242 was derived has an insert size of 1.4Kb. When this clone insert was hybridised to the chromosome 11 YAC panel, several bands were seen in overlapping YACs in the contig. A band was visible in the chromosome 11 only hybrid J1CL4 against a background smear and several bands were visible in total human genomic DNA again against a strong background smear (Figure 4.4). This initially seemed to indicate that this clone, although probably containing a repeat element, (as indicated by the background smear seen in total human DNA) did map to the YAC contig. The pattern of banding in the YACs suggested that this cDNA clone may represent a large gene, exons of which extended across much of the YAC contig. (Figure 4.5 indicates where these exons could lie according to common bands seen in overlapping YACs. Chimaeric YACs D165 and H515 are ignored). The YAC D0485 which spans the translocation breakpoint on chromosome 11 also appears to contain a band, implying that this postulated gene may be in close proximity to the chromosome 11 breakpoint and perhaps even span it. This band was not in common with any YACs overlapping with D0485.

Southern blot hybridisation analysis on the somatic cell hybrid panel using the clone insert as a probe, showed high background smearing despite pre-blocking of the probe with sonicated total human genomic DNA to compete out any repetitive sequences. Since this suggested that the probe contained a substantial repetitive element, primers were designed to the EST sequence of Z39242 in order to reassess the analysis of the YAC panel with a repeat free probe. The cDNA clone corresponding to the EST Z39242 was also sequenced in order to determine the repeat content. The walk primers used to sequence this clone are listed in Table 4.2.

A

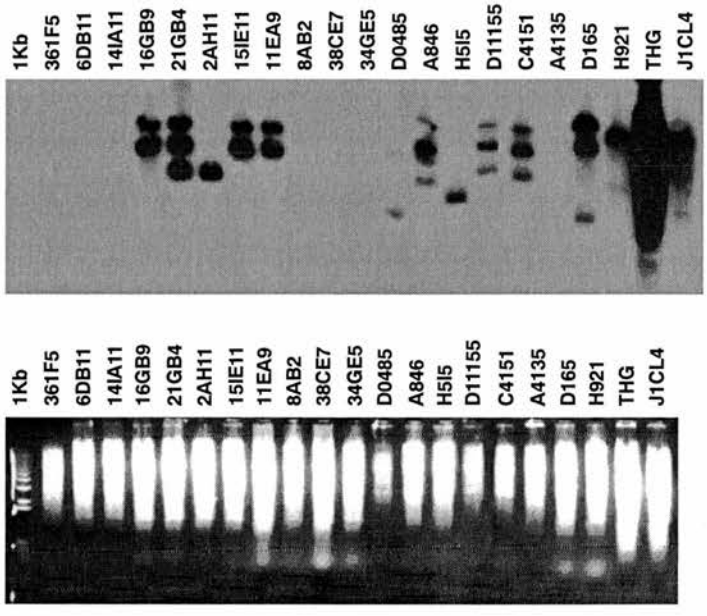


Figure 4.4 A: Southern blot hybridisation analysis of the whole cDNA clone corresponding to the EST Z39242 onto the chromosome 11 YAC panel cut with restriction enzyme *EcoRI*.

B

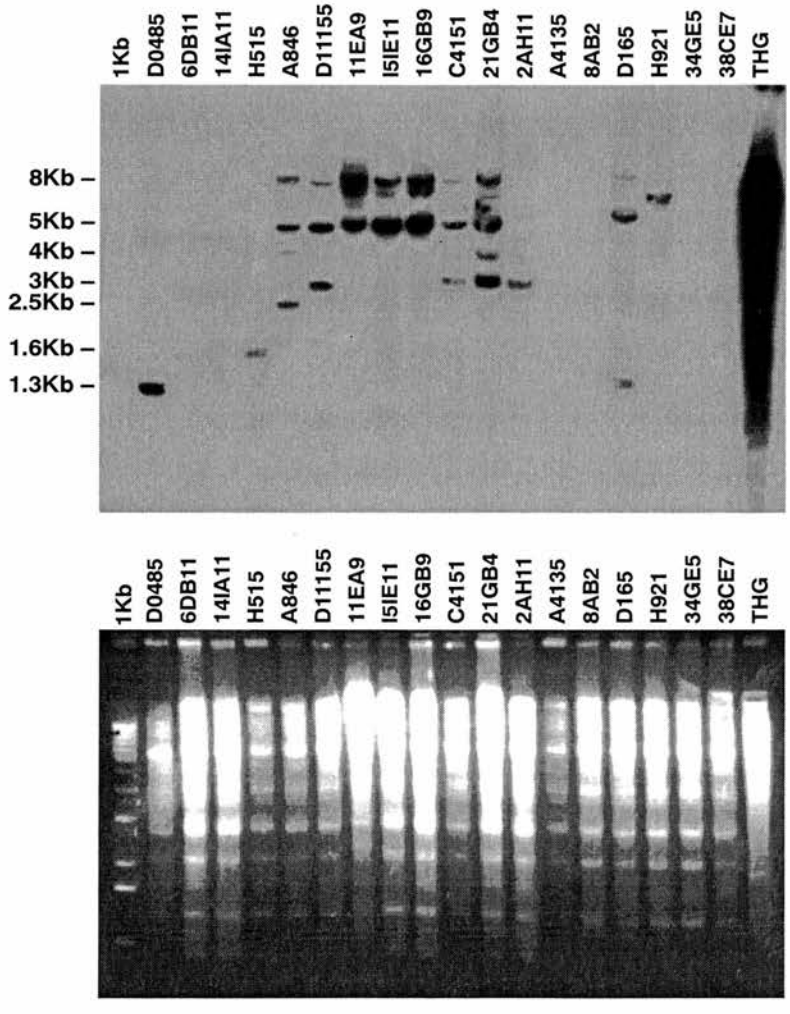


Figure 4.4 B:Southern blot hybridisation analysis of the whole cDNA clone corresponding to the EST Z39242 onto a full length chromosome 11 YAC panel cut with restriction enzyme *EcoRI*

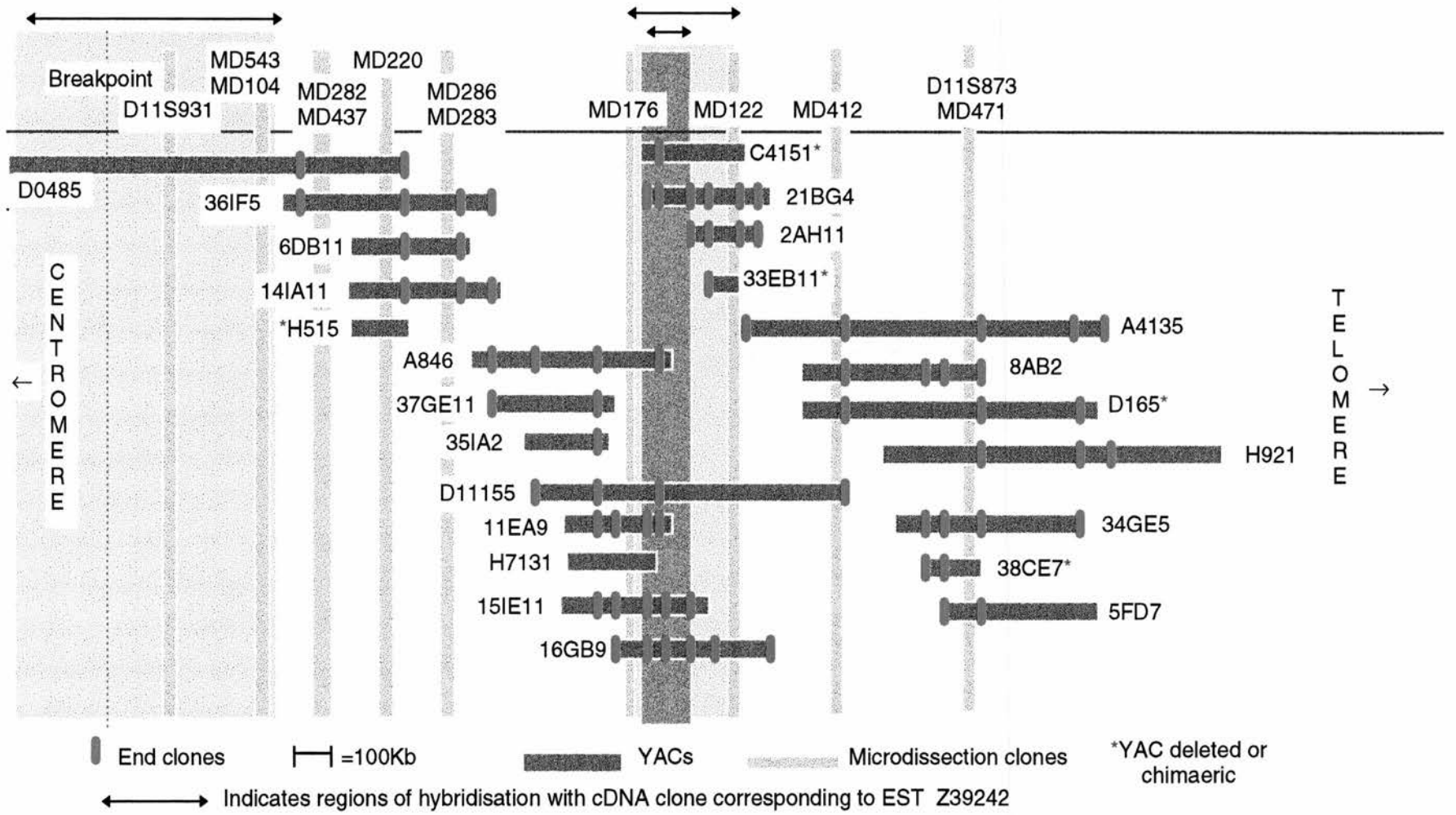


Figure 4.5: Chromosome 11 YAC contiguous clone map (~3Mb) indicating regional hybridisation with cDNA clone corresponding to EST Z39242 (indicated by shaded grey boxes and arrows)

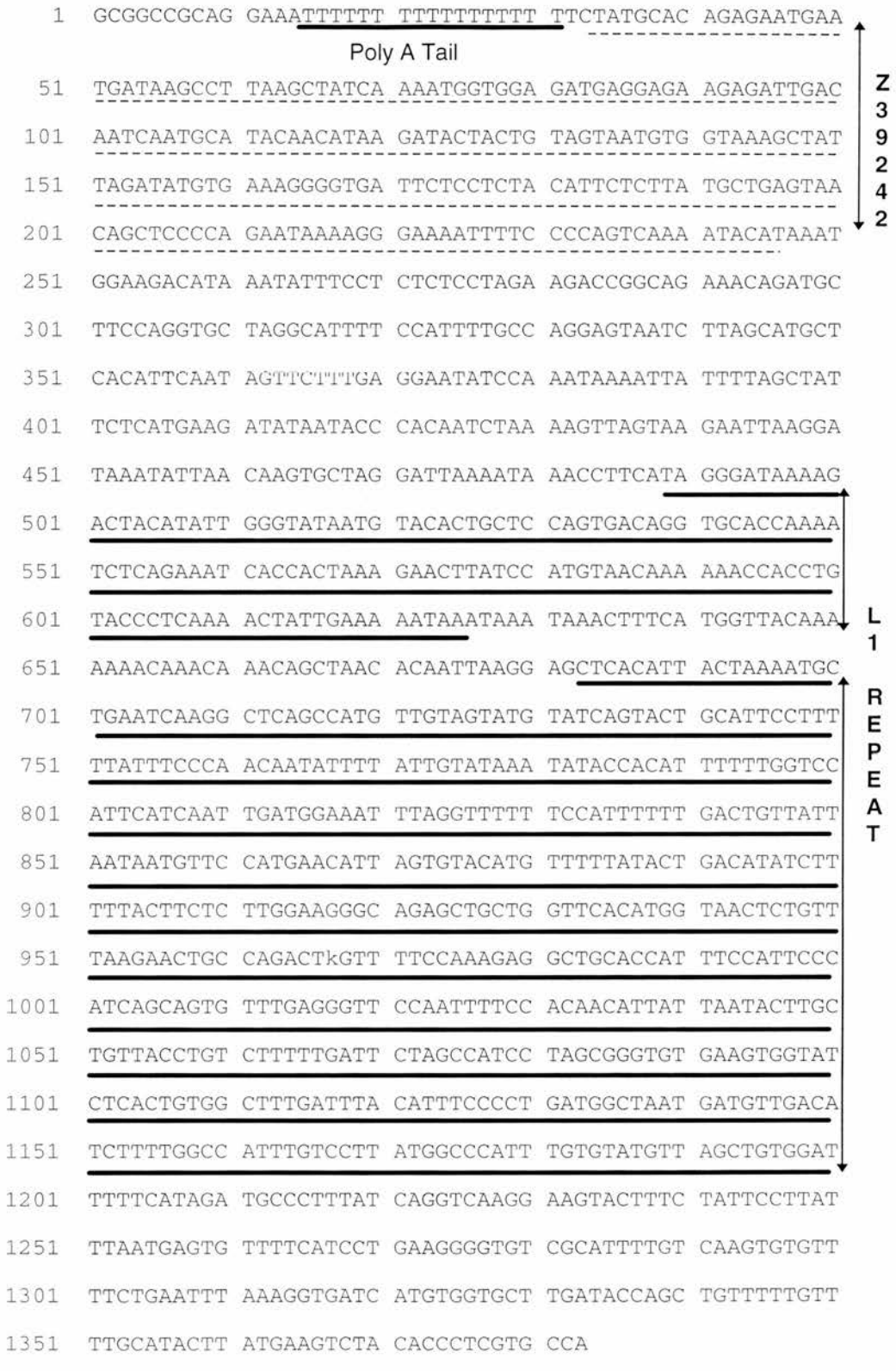
Table 4.2 : Oligonucleotide primers used to sequence cDNA clone corresponding to the EST Z39242. *Primers used in PCR to produce a repeat free probe.

Primer Name	Primer Sequence (5'→3')
Vector Primer M13 (forward)	TGT AAA ACG ACG GCC AGT
Vector Primer M13 (reverse)	CAG GAA ACA GCT ATG ACC AT
*K290	GCA CAG AGA ATG AAT GAT AAG
K97	TGA CAG GGG TGA TTC TCC
K229	CAC CTG GAA GCA TCT GTT
*K872	CTT TTA GAT TGT GGG TAT TAT
K204	CTC AGA AAT CAC CAC TAA
K301	CTG CCA TAC TTG TTT TCC
K568	ATG GGA ATG GAA ATG GTG
K85	TAG CGG GTG TGA AGT GGT ATC
K289	ACT TCA TAA GTA TGC AAA ACA

The sequence of the cDNA clone is shown in Figure 4.6. There is no open reading frame in this sequence and the poly A tail is marked indicating that this clone is derived from the 3' end of the gene.

Repeat database searches (Pythia program Jurka et al 1992) indicated that a repetitive sequence was present and BLAST database searches indicated that several regions in the sequence showed homology to members of the LINE 1 repeat family, which are found at high frequency in the human genome at approximately every 70Kb (Arveiler and Porteous 1992). The fact that there is a large repetitive element in the sequence and no open reading frame implies that this 1.4Kb clone is all 3' untranslated sequence. Southern Blot hybridisation analysis using a Line 1 repeat probe was done on selected YACs from the chromosome 11contig. The result of this hybridisation is shown in Figure 4.7 (This work was carried out by John Maule who also kindly provided the Southern blot hybridisation shown in Figure 4.7). The LINE1 probe ESL1.7 was used to probe the panel of selected chromosome 11 YACs which were cut with restriction enzymes *BssH11* and *Eag1*.

Figure 4.6: Sequence of cDNA clone from which the EST Z39242 is derived. The EST sequence is underlined with dotted line and the repetitive element in the sequence is underlined with bold solid line.



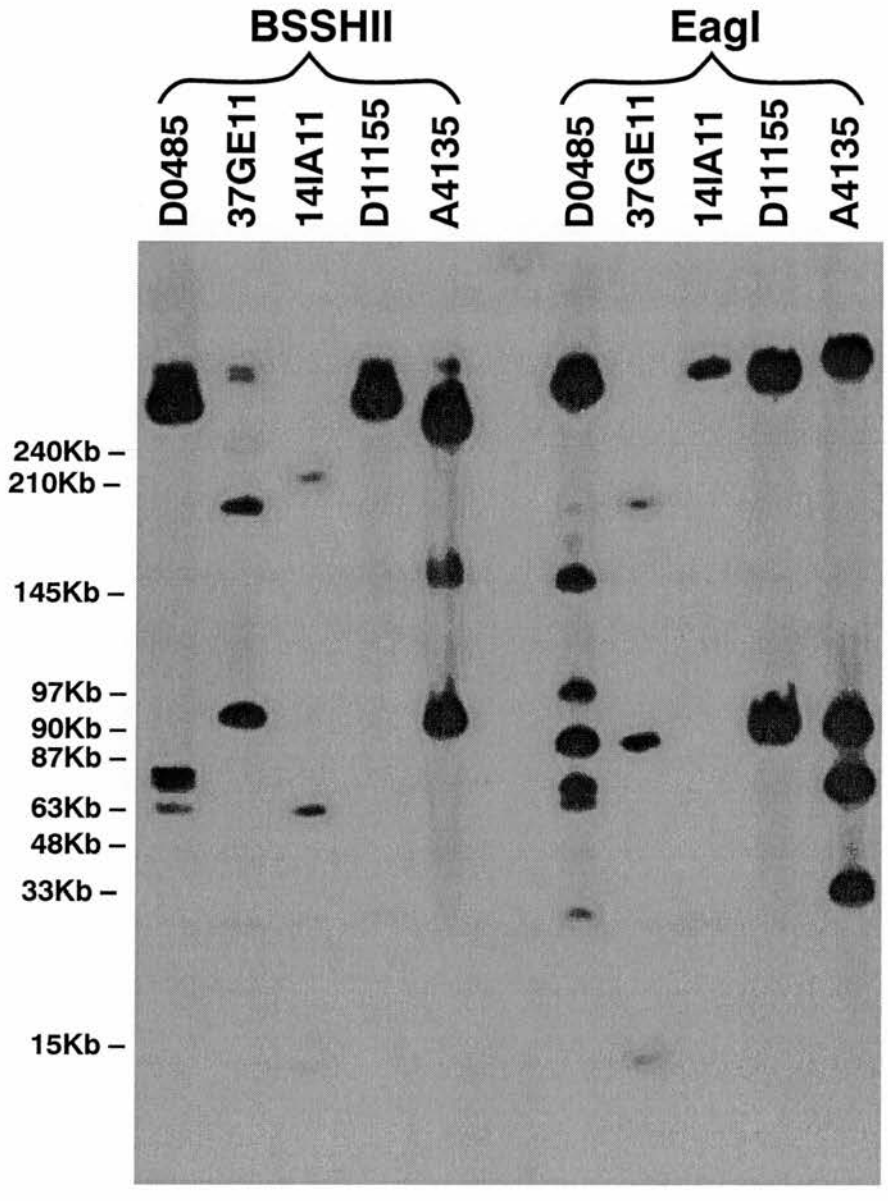


Figure 4.7: Southern blot hybridisation analysis of LINE 1 repeat probe (ESL 1.7) hybridised to selected YACs from the chromosome 11 YAC panel. The YACs have been cut with two restriction enzymes *BssH II* and *Eag I.*
(Southern blot courtesy of John Maule).

The results from this hybridisation showed that only a few bands were seen in the YACs and that this was reminiscent of that seen in the Southern blot hybridisation analysis of the EST Z39242 cDNA clone, in that only a small number of bands were seen in the YACs (although the YACs were cut with restriction enzyme *HindIII* in this panel). This implied that perhaps the bands seen in the initial hybridisation analysis of the EST Z39242 cDNA clone were a result of the LINE 1 repetitive element and that the non repetitive segment of the clone may in fact not map to the YAC contig.
PCR analysis was carried out on the chromosome 11 YACs using primers K290 and 872 (see Table 4.2), in order determine if the non-repetitive component of the cDNA clone did map to the YAC contig. PCR was also done on the cDNA clone DNA (isolated using Qiagen miniprep kit) to prepare a probe for Southern blot hybridisation analysis on the chromosome 11 YAC and somatic cell hybrid panels. PCR conditions were as described in Table 4.3 and 200ng of DNA was used in each reaction. The cDNA clone DNA was used as a positive control along with J1CL4 and total human genomic DNA.

Table 4.3: PCR conditions for amplification of a repeat free section of cDNA clone (from which EST Z39242 was derived) from the chromosome 11 YACs and somatic cell hybrids.

	Denature	Annealing	Extension
Cycle 1	94^{0}C/30sec	55^{0}C/30sec	72^{0}C/25sec
Cycle 2-39	94^{0}C/15sec	55^{0}C/30sec	72^{0}C/25sec
Cycle 40	94^{0}C/15sec	55^{0}C/30sec	72^{0}C/45sec

The results from the PCR indicated that the repeat free segment of the cDNA clone was not contained in any of the chromosome 11 YACs (Figure 4.8). A band of the correct size was however seen in total human genomic DNA and the chromosome 11 only hybrid J1CL4 but not in the mouse or hamster cell lines. It is likely therefore that the bands seen in the original probing using the whole cDNA clone as a probe in Southern blot hybridisation analysis were the result of the repetitive element in the cDNA clone and not a result of the clone mapping to within the YAC contig. Southern blot hybridisation of the PCR product derived from the cDNA clone confirmed that the clone was not contained in any of the chromosome 11 YACs (Figure 4.3). A single band was seen in both total human genomic DNA and in the

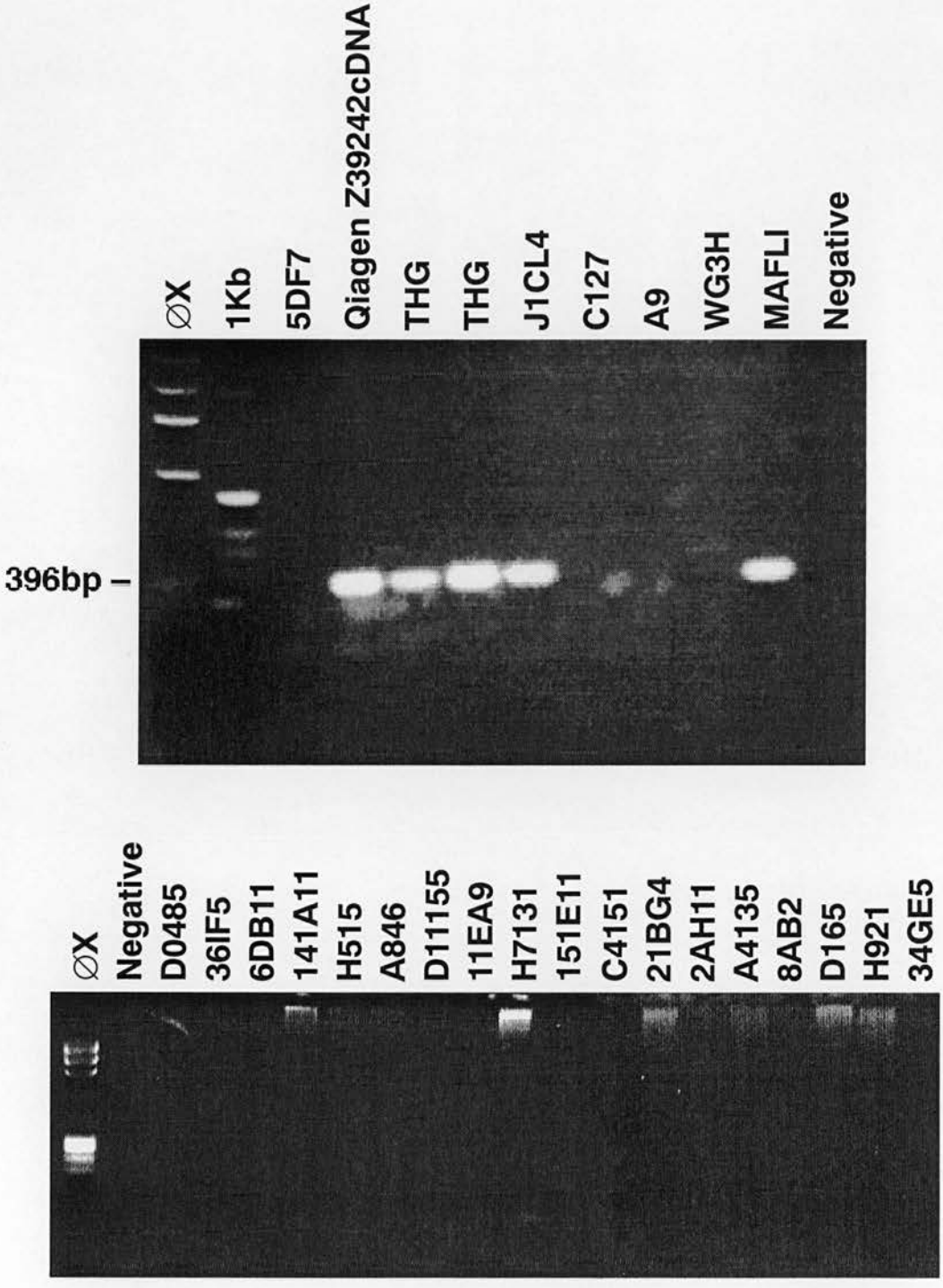


Figure 4.8: PCR analysis showing amplification of the 3' end of the cDNA clone corresponding to EST Z39242 (using primers K290 and 872) from somatic cell hybrids and chromosome 11 YACs.

chromosome 11 only hybrid J1CL4. The band seen in total human genomic DNA was still contained within a background smear but this is most likely due to the higher complexity of the DNA in comparison to that of J1CL4. The band in J1CL4 is much weaker than that seen in total genomic DNA and is a reflection of the lower copy number in this hybrid. A constant band which appears in all the YACs but which is of different size to that seen in the total human genomic DNA and chromosome 11 only hybrid J1CL4 is again seen in this probing although it is very faint. This band is not seen in the yeast only hybrid LUCY2 and so is unlikely to represent a yeast specific fragment, neither does it correspond to the size expected from the Lafmid BA vector sequence (see section 4.2.1).

4.2.2 High resolution mapping of chromosome 11 ESTs by PCR Analysis

The remaining 8 ESTs, for which no cDNA clone could be obtained, were mapped to the chromosome 11 YAC panel by PCR analysis. Primers were designed to the published EST sequences either using the Oligo 4 computer package (Hybaid) or the Primer design program "Primer" at HGMP Resource Centre, Hinxton, Cambridge ("Pimer" - version 0.5 Lincoln et al 1991 (unpublished), MIT centre for genome research and Whitehead institute for biomedical research, Nine Cambridge centre, Cambridge, Massachusetts). Table 4.4 lists the primers and PCR conditions used in the PCR assays. PCR analysis was also carried out on 3 of the ESTs which were originally mapped using cDNA probe hybridisations in order to confirm these results (Z42258, Z42924 and Z46134).

Results from the PCR indicated that none of the ESTs mapped to within the chromosome 11 YAC contig (see Figure 4.9 for examples of PCR results), placing them at a distance greater than 2Mb distal to the breakpoint or greater than 800Kb proximal to the breakpoint. Positive controls J1CL4, total human genomic DNA and negative control LUCY2 Yeast only hybrid, were included in each EST PCR. A negative result in the YACs was only judged to be real if positive bands of the correct size were seen in the positive controls J1CL4 and total genomic DNA and if a negative result was obtained in the yeast only hybrid LUCY2. Several chromosome 11 YACs contained bands in the PCR for EST F00755 but this band was of greater than expected size and corresponded to the size of band seen in the LUCY2 yeast only hybrid. A band of the correct size was seen in the positive control samples and in other chromosome 11 hybrids indicating that this band was derived

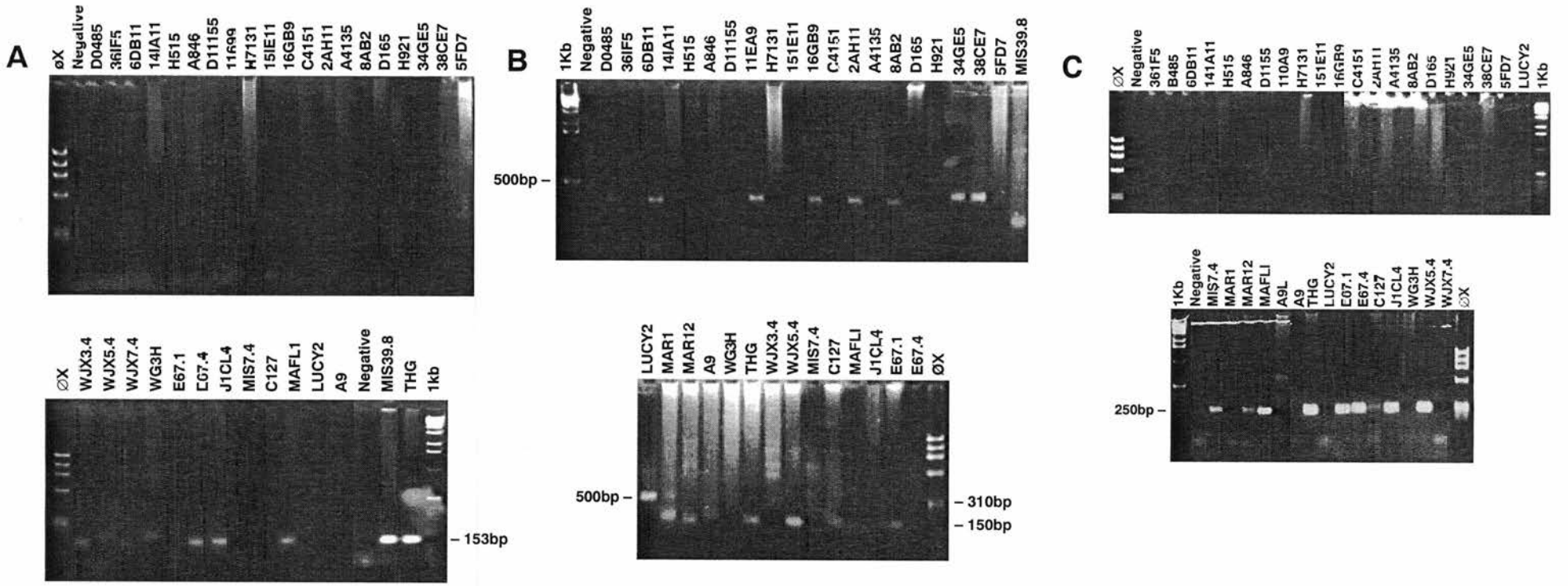


Figure 4.9: Examples of PCR analysis showing amplification of ESTs from chromosome 11 YACs and somatic cell hybrids:
a) PCR analysis for EST F00203 in chromosome 11 YACs and somatic cell hybrids.
b) PCR analysis for EST F00755 in chromosome 11 YACs and somatic cell hybrids. A mouse band (C127) which is the same size as that seen in the human lane is present and excludes the possibility of mapping this EST with the somatic cell hybrids (see text for further explanation). A larger sizeyeast specific band is also visible in the chromosome 11 YACs and in the Yeast only hybrid LUCY2.
c) PCR analysis for EST Z28653 in the chromosome 11 YACs and somatic cell hybrids.

from yeast and probably represents a homolog of this EST in yeast (see Figure 4.9b).

PCR analysis was also done on the somatic cell hybrids to confirm the position of these ESTs as either proximal or distal to the breakpoint and to refine their mapping positions further. Results from this analysis are summarised in Table 4.5.

Table 4.4: Oligonucleotide primers and conditions used to amplify EST sequences from chromosome 11 YACs.

EST (Genbank I.D)	Primers (5'→ 3')	Annealing Temperature	Product Size (bp)
Z19327	AGG CAC AAG TTT AAG CAG CAT TGC CAG TCT GTA TTG AGG GGT	55°C	159
Z28525	AGT CTC CAC AAA TGG TCT GCT AAT GCC GTA TGT TGT AGG GGT	55°C	167
Z28611	GTC TGT CTC TGC GGC TTC TGT CCT TCA AGT TCC AAG CGT CT	N/D	250
Z28653	TTG AAA ACA CTG AAA ACC TCT TGA ACT GAA CCT TGG ACT GAT	50°C	250
F00203	AGT GGT CGC CAA GTA AGA GG CTT GGC ATG TGA GGG CTA TT	57°C	153
F00755	GTG GAG CAC ATT ATA GAC AG AAT GAG TTG TTG TTG TTG AT	50°C	155
F00845	GTG GTT ACA TCT ATT CTT GCT TGC CTT GGA TAC TTG GCA TTT	50°C	100
Z45606	GAA ATC TGT CCC GTA AAG ATA AGA CAC CAG CAG TAG	50-40°C (Touchdown)	96
Z42258	ATG TTC TTT TCA GCG GTC AAA TAG GGT TGT AAA AGG	45°C	270
Z42924	TAA CCT ATT TCC ACA TTC TGT TAA AGA TGA CTG TGA	45°C	300
Z46134	ATG TTT ACC TCT TTC TTT GAA AAG ATA CTC AAGGTT	45°C	150

Table 4.5 : Summary of results from PCR analysis of chromosome 11 ESTs on somatic cell hybrids.

Hybrid	EST							
	Z19327	Z28525	Z28653	F00203	F00755	F00845	Z45606	Z46134
THG	+	+	+	+	+	+	+	+
J1CL4	+	+	+	+(+>)	+	+	+	+
MIS39	+	-(+>)	N/D	+	N/D	-	-	-
MIS7.4	+	+	N/D	-	N/D	+	+	+
MAR1	+	N/D	-	N/D	+	N/D	N/D	-
MAR12	+	N/D	+	N/D	+	N/D	N/D	+
WJX3.4	-	-	-	+	-	-	-	-
WJX5.4	-	+	+	-(+>)	+	+	+	+
WJX7.4	-	-	-	+	-	-	-	-
E67.1	+	-	+	-	+	+	+	-
E67.4	+	-(+>)	+	+	+	+	+	-
RAG	+	N/D	N/D	N/D	N/D	N/D	-	N/D
C127	+	-(+>)	-	-	+	-	-	-
WG3H	-	-	-	-(+>)	-	-	-	-.
A9	+	-	-	-	+	-	-	-
LUCY2	-	-	-	-	+(+>)	-	-	-
MAP REGION	?	DER1 D11S351-THY1 11q22.3	DER1	DER11 D11S533-OMP 11q13-q14	?	DER1	DER1	DER1 D11S351-THY1 11q22.3

THG - Total human genomic DNA
\+ - PCR product of expected size present in somatic cell hybrid
\- - No PCR product of the expected size was seen in somatic cell hybrid
N/D - Hybrid not done
(+>) - Band of greater than expected size present.
? - Map location undetermined.
DER1- derived chromosome 1 indicating that this EST maps to distal the breakpoint on chromosome 11.
DER11- derived chromosome 11 indicating that this EST maps to proximal the breakpoint on chromosome 11.

Two of the ESTs (Z19327 and F00755) could not be localised to either above or below the translocation breakpoint on chromosome 11 as a similar size band to that expected for the EST was seen in the mouse only hybrid, thus masking the status of all the hybrids in a mouse background (including the MIS and MAR hybrids). In EST Z42258 (not shown in Table 4.5) the mouse only hybrid similarly contained a band of equivalent size to that expected from the EST in human DNA which prevented it being further mapped.

It would be possible to map these EST further by determining restriction enzyme sites which were different between the two homologs (mouse and human). Restriction enzyme digests could then be carried out on the PCR products, allowing differentiation between mouse and human bands and those hybrids containing the human EST to be determined. The PCR on EST Z28525 also showed a band in the mouse only hybrid but this band was distinguishable from the expected human product and therefore did not pose any problem. One ESTs, F00203, had a product present in the hamster only hybrid but this again was of a different size to the expected human product and was distinguishable from it.

The ESTs Z45606, F00845, Z28653 were unable to be more finely mapped, other than to lying distal to the translocation breakpoint as determined by the MIS and MAR hybrids. The inability to more finely map these ESTs was a result of them all showing the same positive pattern in the hybrids, namely MIS 7.4, WJX5.4, E67.1 and E67.4 for which no corresponding region existed from previous characterisation of these hybrids with chromosome 11 markers (no region distal to the breakpoint is positive for both E67.1 and E67.4, see Table 3.3 and 4.5). This is likely to be a result of either cross contamination of the two hybrids with each other so that they both appear positive in PCR analysis or undetected positive regions of chromosome 11 in the hybrids due to the large distances between the markers used to characterise them. It is possible that a region which is positive for both E67.1 and E67.4 is present distal to the breakpoint but more markers in the intervening distance would be required to further characterise the hybrids.

The Human Transcript Map which has been produced by Schuler et al (1996) (and is continually being added to) has mapped many of the ESTs described in the Rosier et al (1995) paper subsequently to this work being undertaken. The results of this mapping are shown in Table 4.5.

Since the ESTs, which could not be mapped due to mouse or hamster homologs masking the PCR results, have now been mapped on the Human Transcript Map (Schuler et al 1996) (see Table 4.6) it was unnecessary to repeat the mapping by restriction enzyme digestion of the somatic cell hybrid PCR products. One EST, Z28611, was refractory to PCR and could not be amplified despite using several differing annealing temperatures and magnesium concentrations. A new set of oligonucleotide primers was designed to the EST sequence but did not improve the PCR. However, this EST has now been mapped by the Human Transcript Map and does not lie in the vicinity of the chromosome 11 breakpoint.

Since none of the ESTs appear to map to the region around the translocation breakpoint they are deemed to be too far from the breakpoint to be further pursued as candidate genes involved in the psychiatric diagnosis in the translocation family K26.

Discussion:

High resolution mapping of Expressed Sequence Tagged Sites (ESTs) derived from sequencing the ends of cDNA clones is a rapid method by which to identify candidate gene transcripts from a region which has been genetically linked to a disease. There are various pitfalls to using these however such as low copy repetitive elements which may be present in the corresponding cDNA clones which may give somewhat misleading results as has been demonstrated in the current study. This is often due to the presence of 3' untranslated regions of the gene being present in the clones which are more likely to contain repetitive elements than coding regions. Since the ESTs are often derived from cDNA libraries which have been oligo dT primed and therefore select the 3' most end of the gene by virtue of the poly A tail it is likely that many cDNA clones corresponding to ESTs will contain repetitive elements and this should be borne in mind when using these clones.

Of the 20 ESTs which were mapped to the region around the translocation breakpoint on chromosome 11 by Rosier et al (1995), high resolution mapping showed that none resided within the chromosome 11 YAC contiguous clone map. This placed these ESTs at a distance of greater than 2Mb distal and 800Kb proximal to the chromosome 11 breakpoint. It is therefore unlikely that any of the genes corresponding to these ESTs are involved in the psychiatric diagnosis in the K26 translocation family.

There are many ways that a translocation event may disrupt gene function such as by affecting the chromatin structure around a gene, by disrupting long range control elements of gene function or by creating gene fusion products. There are examples of translocation breakpoints having an effect on a gene some distance away and causing disease. One such example is seen in the case of the PAX6 gene, mutations in which cause the eye disorder Aniridia. Several families have been described where the chromosomal breakpoints have been up to 85Kb distal to the PAX6 gene and caused Aniridia. Since no mutations in the PAX6 gene were seen this was thought likely to be because a chromosomal rearrangement which caused an inappropriate chromatin environment which does not allow normal expression of

the gene (Fantes et al 1995). Another example occurs in Campomelic dysplasia, which is a skeletal malformation syndrome. This disease is caused by mutations in the SOX9 gene on chromosome 17q and translocation families have been described in which the breakpoint is 130Kb from the SOX9 gene (Wirth et al 1996). Translocations events in Burkitt's Lymphoma have shown that when the gene responsible for the condition, c-myc, is brought under the influence of B-cell control elements on chromosomes 2, 14 or 22 the breakpoints can be up to 340Kb away from the c-myc gene and still lead to constitutive up-regulation expression of the gene (Joos et al 1992). Although it may be possible that longer range effects which could act at up to 800Kb from the breakpoint may exist, none have been demonstrated to date. It is therefore unlikely that any of the genes corresponding to these ESTs are involved in the psychiatric diagnosis in the K26 translocation family. Several genes which reside within 800Kb proximal to the translocation breakpoint on chromosome 11 have been identified (Devon et al 1997) as a result of the positional cloning strategy being undertaken to elucidate the genes involved in the psychiatric diagnosis in this translocation family and are currently being persued.

The Human Transcript Map (Schuler et al 1996, World Wide Web site location: http://www.ncbi.nlm.nih.gov/SCIENCE96/. Site hosted by National Centre for Biotechnology Information, National Institute of Health, Bethesda. MD 20894 USA) has recently localised many of the ESTs from around the breakpoint region to an integrated chromosome 11 radiation hybrid, cytogenetic and genetic map. This data has confirmed the location of the ESTs as either above or below the breakpoint. The ESTs whose map positions were able to be determined by the chromosome 11 somatic cell hybrids are slightly different from those of the Human Transcript Map. However, these positions are at the same general location. For example, those ESTs mapping to the interval D11S351 to THY1(corresponding to 110-112cM on Human Transcript Map, cytogenetic location approximately 11q22.3 refer to figure 4.1) such as Z28525 and Z46134 are mapped in the Human Transcript map to region D11S1325-D11S1778 (105-106cM) respectively. This corresponds to a cytogenetic position of 11q21-q22. EST F00203 was mapped to the region between D11S533 and OMP, corresponding to cytogenetic region 11q13.5-q14.1, by somatic cell hybrid PCR analysis. The marker D11S533 is in very close proximity to the marker D11S911 in the Eugene McDermott Centre for Human Growth and Development integrated chromosome 11 map (World Wide Web Site location: http://mcdermott.swmed.edu/. Site hosted by University of Texas South Western

Medical Centre, Dallas, Texas, USA). The Human Transcript Map places this EST at interval D11S916-D11S911 at approximately cytogenetic location 11q13. This EST's map position is very similar to that seen in the Transcript map. As a result of a greater number of STS markers on the Human Transcript map and very detailed characterisation of the radiation hybrids used to localise ESTs it is quite possible that the mapping positions on this map will differ slightly from those of the chromosome 11 somatic cell hybrids. The latter were characterised with far less markers and therefore the cytogenetic location inferred by these results is less specific than that of the Transcript Map. However, given that the ESTs which were mapped by the somatic cell hybrid method correspond reasonably well to the map positions in the Transcript Map the viability of using this method for high resolution mapping is again demonstrated.

The ESTs which have been investigated in this study for involvement in the psychiatric diagnosis in the translocation family have been considered as candidate genes solely on the basis of their position in close proximity to the translocation breakpoint.

None of the ESTs will be considered further at the present time as candidate genes due to their distant location from the translocation breakpoint. However, until the whole gene has been mapped we cannot exclude the possibility that the 5' end is in the region of the breakpoint (since the ESTs are usually designed to the 3' end of the gene and some genes can be extremely large). It is possible that if further biological information becomes available on the genes that these ESTs represent which is of interest, such as further mapping or functional and expression data which may indicate their possible involvement in the psychiatric diagnosis of the K26 pedigree, these genes may be further considered as candidate genes in the absence of any stronger candidate gene residing in closer proximity to the breakpoint. Information on these ESTs, available through various databases and World Wide Web sites, will therefore be monitored.

Table 4.6: ESTs described in Rosier et al (1995) mapping positions in Human Transcript Map (Schuler et al 1996 and on WWW at the National centre for biological information, National Institute of Health website location http://www.ncbi.nlm.nih.gov/SCIENCE96/)

EST (Genbank I.D)	Mapped to Transcript map	Position (Marker and cM location)	Homology
F00504	YES	D11S1778 (106cM) 11q21-q22	mitochondrial acetoacetyl CoA thiolase
F00690	YES	D11S1893-D11S938 (109-110cM) 11q21-q22	Alpha Crystallin B chain
F05464	NO		C Protein (Homo sapiens) (M16342)
F02133	YES	D11S1354-D11S931 (92-96cM) 11q14-q22	NONE
Z38957	YES	D11S1325-D11S1778 (105-106cM) 11q21-q22	NONE
F04857	NO		Zinc finger protein (weak) (X71623)
F02376	YES	D11S938 (110cM) 11q22-23	NONE
Z39212	YES	D11S1354-D11S931 (92-96cM) 11q14-q22	African swine fever virus (very weak)
Z39242	YES	D11S938 (110cM) 11q22-23	NONE
Z42258	YES	D11S1893 (109cM) 11q21-22	NONE
Z42924	NO		NONE
Z46134	YES	D11S1325-D11S1778 (105-106cM) 11q21-q22	NONE
Z19327	YES	D11S1311-D11S923 (97-102cM) 11q14-q22	NONE
Z28525	YES	D11S1325-D11S1778 (105-106cM) 11q21-q22	NONE
Z28611	NO*	D11S1333-D11S917 (98-100cM) 11q14-q22	KIAA0092 gene - related to smooth muscle myosin (D42054)

Z28653	YES	D11S1778 (106cM) 11q21-q22	NONE
F00203	YES	D11S916-D11S911 (80-84cM) 11q13	Thyroid hormone-inducible hepatic protein (Rat) (strong)
F00755	NO		ALU repeat (moderate)
F00845	YES	D11S1778-D11S1893 (106-109cM) 11q21-q22	ALU repeat (moderate)
Z45606	YES	D11S1893-D11S938 (109-110cM) 11q21-q22	NONE

* This EST was not mapped in the Human Transcript map but shows strong homology to the KIAA0092 protein which has been mapped to the region indicated.

Chapter Five

An allelic association study of two polymorphic markers in close proximity to a balanced translocation t(1:11) breakpoint which co-segregates with mental illness.

5.1 Introduction

One of the most popular strategies for finding genes involved in psychiatric disorders and many other complex disorders, has been to carry out genetic linkage analysis on large families with many affected members. However, such studies have failed to identify unequivocally the gene(s) involved in schizophrenia, most likely due to the complex nature of the disorder (polygenic non-Mendelian inheritance, probable aetiologic heterogeneity and difficulties with precise phenotype definition). Recent world-wide collaborations and methodological refinements have enabled large numbers of families to be studied, thus improving the power of the technique and several candidate regions have been identified, most notably chromosome 6p (Wang et al, 1995, Straub et al 1995) and 22q (Pulver et al, 1994; Coon et al, 1994). Further replication and refinement of these regions is needed. Population based association studies, in which the frequency of an allele at the marker locus in disease population is compared to that of an ethnically matched control population, provide an alternative approach to family based linkage analysis and have been widely used in psychiatric genetics.

Association studies have considerable appeal since they can detect genes that contribute to only a relatively small proportion of the overall liability to developing a disorder. It is likely that most psychiatric disorders result from the combined effects of several genes of moderate or small effect and therefore association studies may be a particularly good way of identifying these genes. Association studies are non-parametric and therefore do not rely on any prior assumptions, other than the disorder having a genetic component, unlike conventional genetic linkage studies in which mode of inheritance, penetrance levels etc. must be correctly specified. Another distinct advantage of association studies is that large affected families, which are often difficult to obtain, are not required.

Association studies have been successfully performed in a number of complex disorders such as late onset Alzheimer's disease with apolipoprotein E 4 allele (Corder et al 1994) and myocardial infarction with angiotensin converting enzyme (Cambien et al 1992). A recent and impressive demonstration of the power of association studies to detect the subtle effects of genes on multifactorial quantitative traits, is the association between a dopamine D4 receptor gene polymorphism and the personality trait of novelty seeking behaviour (Ebstein et al 1996). The initial report of association in a group of 124 Israeli subjects whose

personality scores were based on tridimensional personality questionnaire (TPQ) has been replicated using a different personality questionnaire in an ethnically different population (Benjamin et al 1996) and family studies have demonstrated that the effect is genetic and not due to population stratification. Association studies are being widely used in attempts to detect quantitative trait loci involved in personality as they provide the necessary statistical power to detect relatively small gene effects that contribute to complex behavioural traits.

The positional cloning strategy which is currently being employed on the K26 translocation family to elucidate the relationship between the translocation event and the psychiatric diagnosis in this family is based on the hypothesis that a gene or genes involved in mental illness in this family reside at or near the translocation breakpoint on chromosome 1 and/or 11. The regions around the chromosome 1 and 11 breakpoints have been cloned as described in sections 1.16 and 7.2. Two polymorphic markers have been identified as lying in close proximity to the translocation breakpoint, as a result of this positional cloning strategy. A 2.7Kb derived chromosome 1 fragment was isolated from hybridisation of a genomic library made from a translocation patient DNA using a 2.15Kb cosmid derived probe which spanned the wild-type chromosome 11 breakpoint. This derived chromosome 1 probe was used as a probe back onto the genomic translocation patient library and identified a 7Kb wild-type chromosome 1 fragment. This wild type chromosome 1 fragment was sequenced in order to determine differences between the wild-type and derived chromosomes in this region (see section 7.2). As a result of this sequencing and subsequent database searching, a polymorphic triplet repeat, which showed strong homology to the STS marker D1S1621, was identified. This triplet repeat was approximately 300bp distal to the chromosome 1 breakpoint.
The dinucleotide repeat marker, D11S931, is in close proximity to the chromosome 11 breakpoint and resides at approximately 85Kb proximal to the translocation breakpoint.
The evidence of linkage between chromosome 1 and 11 and mental illness is sparse with few independent reported genetic linkage results from around the region of the translocation breakpoint. Genetic linkage studies by our group using chromosome 1 and 11 markers in close proximity to the breakpoint have provided no further evidence of linkage (Muir et al in preparation). However, given that association studies are a sensitive method for detecting genes of moderate to weak

effect, markers D1S1621 and D11S931 provide an opportunity to investigate an association between the translocation breakpoint region and a postulated schizophrenia susceptibility gene residing in close proximity to the breakpoint, in a random population of affected subjects versus normal controls. These subjects were selected from the same geographical region as that of the translocation family.

5.2 Pilot study to assess the polymorphism of marker D1S1621

A pilot study was carried out using the marker D1S1621 in order to determine how polymorphic this marker was and if it would be sufficiently informative, in terms of its polymorphism, to be used in the association study. Several members of the translocation family were genotyped as described above for marker D1S1621. It was apparent that five alleles were present, with an excess of one of these alleles being seen in most of the family members (allele 261bp). When a larger population was studied (in the association study described below) a sixth rarer allele was also identified (267bp). The marker D1S1621 was deemed polymorphic enough to be used in the study based on the fact that it had six alleles which seemed to all (with the exception of allele 6, 267bp) appear frequently in the population studied.

Based on the information obtained from the allelic association study on the control sample (described below), the heterozygosity of the marker D1S1621 was assessed.

Heterozygosity can be determined by the equation:

$$HET = 1 - \sum_{i=1}^{n} P_i^2$$

n = number of alleles
Pi = population relative frequency of the i^{th} allele

HET - Probability that a random individual is heterozygous for any two alleles at a gene locus.

For marker D1S1621 the heterozygosity value is 0.77 as determined by the UNILINK computer program (Terwilliger and Ott 1994).

5.3 An allelic association study of D1S1621 and D11S931 in schizophrenia subjects, unipolar depressed subjects compared to control subjects.

The frequencies of the alleles of D11S1621 and D11S931 were studied in a random population of subjects with schizophrenia and also in a random population of unipolar depression subjects (a diagnosis also prominent in the t(1:11) family) and compared to ethnically matched control subjects in order to determine if there was a difference in allele frequencies between the affected populations and the control population.

5.3.1 Subjects:

The schizophrenia subjects (n=105) and the unipolar depressed subjects (n=84) were all either in-patients or out-patients at the Royal Edinburgh Hospital, Morningside Park, Edinburgh. All patients were interviewed by psychiatrists using the Schedule for Affective Disorder and Schizophrenia-lifetime version (SADS-L) (Endicott and Spitzer 1978). Diagnoses were based on RDC and DSM IV criteria (Spitzer et al 1978 and American Psychiatric Association 1994). Control subjects were obtained from the general population and were excluded from the study if they had a personal or family history of psychiatric illness. All subjects originated from the south east of Scotland and gave informed consent to take part in this study.
Standard procedures were used to extract DNA from peripheral blood samples.
Subjects were genotyped for markers D11S931 and D1S1621 by the polymerase chain reaction (PCR) using a fluorescent primer and electrophoretic analysis was carried out using a 40-lane automated laser fluorescence (ALF) sequencer.

5.3.2 Genotyping subjects for markers D11S931 and D1S1621

Amplification of D1S1621 marker was achieved using primers 5'-TTTCTCACCTTTAAATGTCATCA-3' and 5'-CCAGTACGCAGATGGTCCTA-3'. The first of these primers was used as described from information obtained from the GENBANK database and Whitehead institute/MIT centre for genome research databases (http://www.genome.wi.mit.edu/)(Hudson et al 1995). The second of these primers was designed to the sequence extending just beyond the marker D1S1621 corresponding to the sequence obtained from the 7Kb wild-type

chromosome 1 sequence around the breakpoint as the primer described in the database for this marker was refractory to PCR. This PCR produced a product of between 252bp and 267bp. D11S931 was amplified with oligonucleotide primers 5'-ATGTTGGTAGGTATTCT- 3' and 5'-GAGAAATAGTATGTGTTTGCC -3' which were designed to the EST sequence described in the database in order to obtain a product which was of a different size to that for D1S1621, since the primers described in the database would have produced a product of approximately 250bp. The difference in PCR product size was desirable to allow both the PCR products, from the same patient DNA, to be run in the same lane of the polyacrylamide gel, thus reducing the number of gels required for the genotyping.

Tables 5.1 and 5.2 show the PCR conditions for markers D1S1621 and D11931 respectively.

Table 5.1 PCR conditions for marker **D1S1621**:

	Denaturation	Annealing	Extension
Cycle 1-27	94°C/30sec	55 °C/75sec	72°C/15sec
Cycle 28	94°C/30sec	55°C/75sec	72°C/6mins

Table 5.2 PCR conditions for marker **D11S931**:

	Denaturation	Annealing	Extension
Cycle1	94°C/5min	47°C /30sec	72°C /1min
Cycle 2-29	94°C/1min	47°C /30sec	72°C /1min
Cycle 30	94°C/1min	47°C /30sec	72°C /2min

Electrophoresis was carried out on a 40-lane automated laser fluorescence (ALF) sequencer (Pharmacia).

After the PCR reaction, the products from both reactions (D11S931 and D1S1621), for the same patient, were run in the same lane on the ALF sequencer (since the products are of different size) with appropriate size markers and gold standards (kindly provided by Stewart Morris). Automated genotyping was carried out using 'ALF' manager and 'ALP' software (He *et al* 1995, Mansfield *et al* 1994). All gel results were also checked manually.

D1S1621:

For marker D1S1621, 142 control subjects, 105 schizophrenia subjects and 84 unipolar depressed subjects were genotyped. The PCR detected 6 alleles of sizes 252bp, 255bp, 258bp, 261bp, 264bp and 267bp. The allele 258bp was the most frequently seen allele in the control population, occurring in approximately 33% of the controls tested. The 267bp allele was seen more rarely than the other alleles occurring in only 2% of the control population. The most common genotype seen in the control population was 3/3 (258bp) (14%) with 2/3 (255bp/258bp) (13%) also being seen at high frequency. No genotype 6/6 (267bp) was seen in any of the groups tested.

D11S931:

In the case of marker D11S931, 143 control subjects, 100 schizophrenia subjects and 84 unipolar depressed subjects were genotyped. The marker D11S931 has 4 alleles of sizes 149bp, 159bp, 163bp, 165bp. The 165bp allele was the most commonly seen in the control population (37%) and the 159bp allele is the least common (10%). The most common genotype for marker D11S931was 1/4 (149bp/165bp) (24%) and the least common 2/2 (159bp/159bp) (1%).

Figures 5.1 And 5.2 are graphical representations from the ALF automated sequencing gel which shows the alleles of markers D1S1621 (Figure 5.1) and D11S931 (Figure 5.2).

Figures 5.3 to 5.6 are graphical representations of either percentage allele frequencies or percentage genotype frequencies.

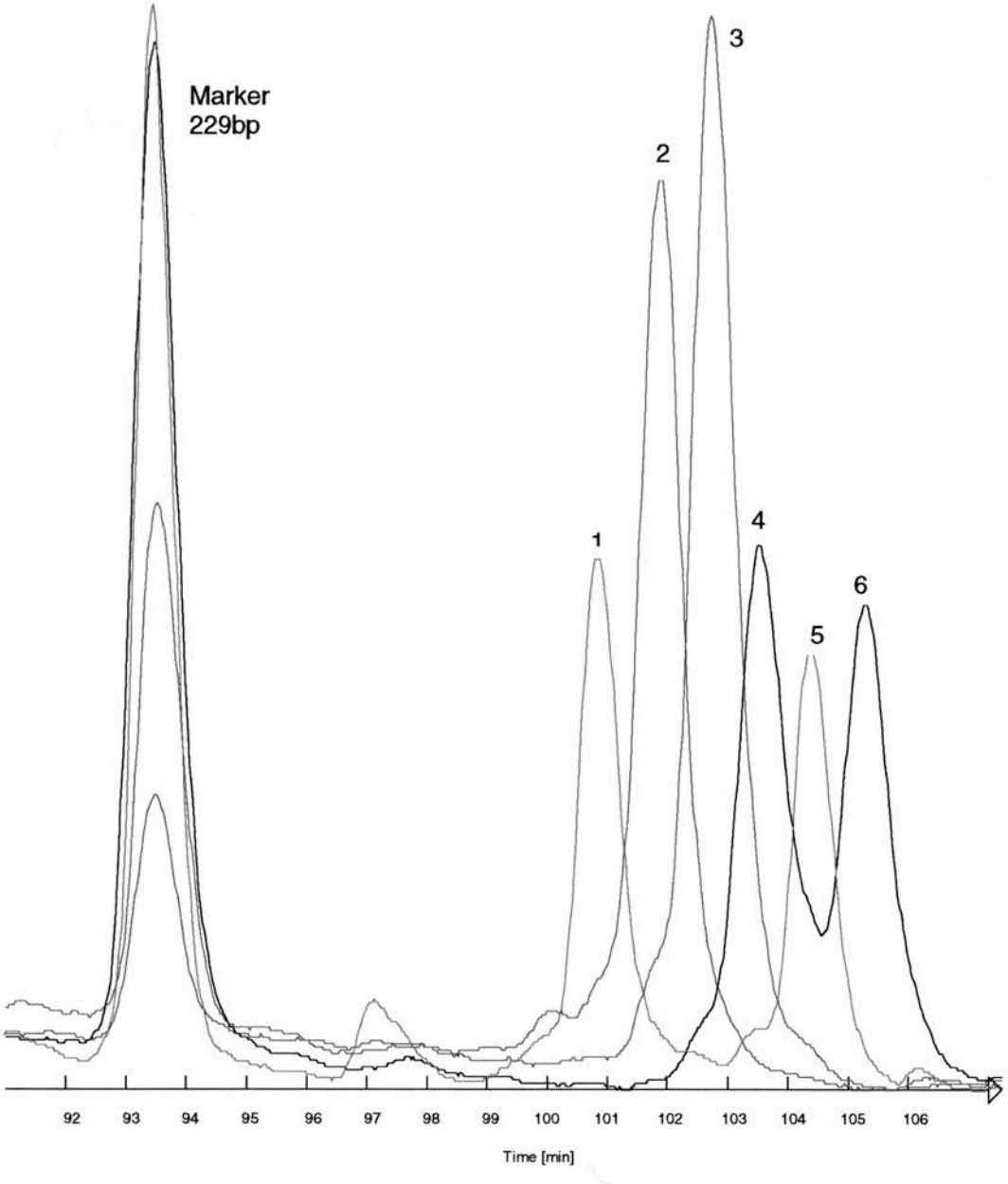


Figure 5.1: Graphical representation of ALF sequencing gel showing the six alleles (1=252bp, 2=255bp, 3= 258bp, 4=261bp, 5=264bp and 6=267bp) for marker D11S931 along side a size marker (229bp). The peak height is representative of the intensity of the fluorescent band seen on the sequencing gel and the X axis denotes the time scale (minutes). Each colour represents a different patient DNA so that the genotype of the patient represented in blue is homozygous 3/3 (corresponding to the allele at 258bp).

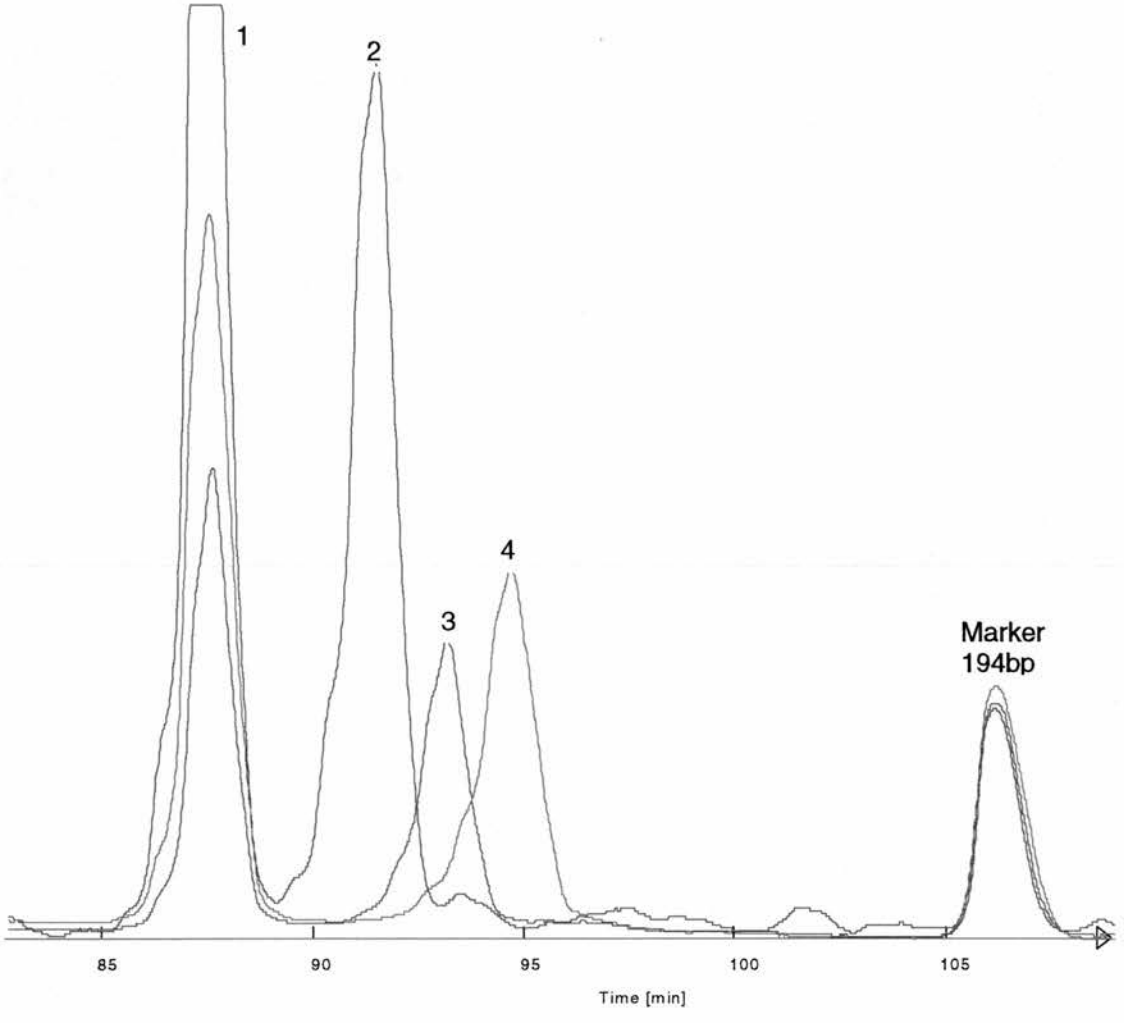


Figure 5.2: Graphical representation of ALF sequencing gel showing the four alleles (1=149bp, 2=159bp, 3=163bp and 4=165bp) for marker D11S931 along side a size marker (194bp). The peak height is representative of the intensity of the fluorescent band seen on the sequencing gel and the X axis denotes the time scale (minutes) over which the gel was run. Each colour represents a different patient DNA so that the genotype of the patient represented in blue is heterozygous 1/3 (corresponding to alleles at 149bp and 163bp).

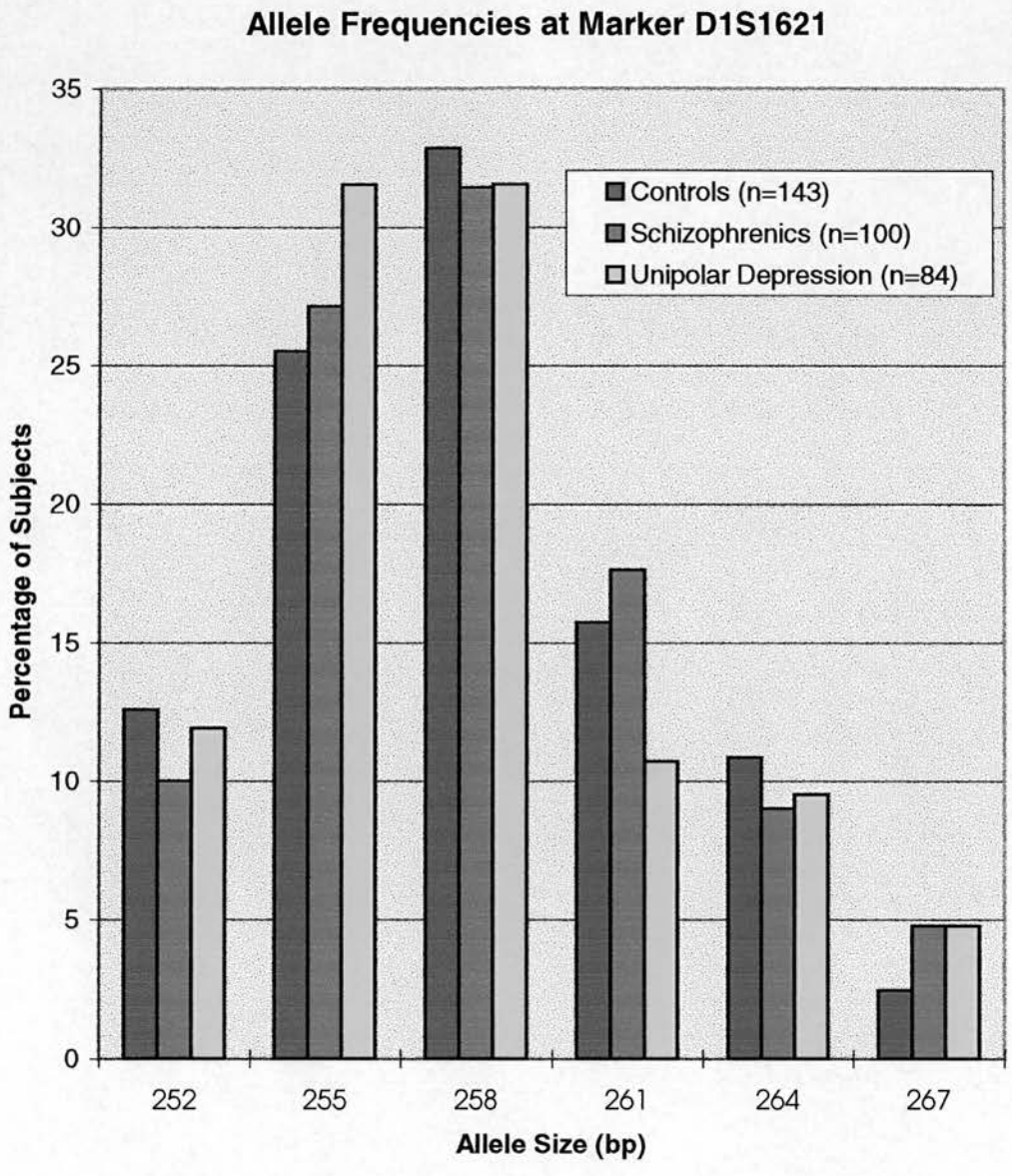


Figure 5.3 Graphical representation of percentage allele frequencies at marker D1S1621 for the schizophrenia subjects, unipolar depressed subjects and the control subjects.

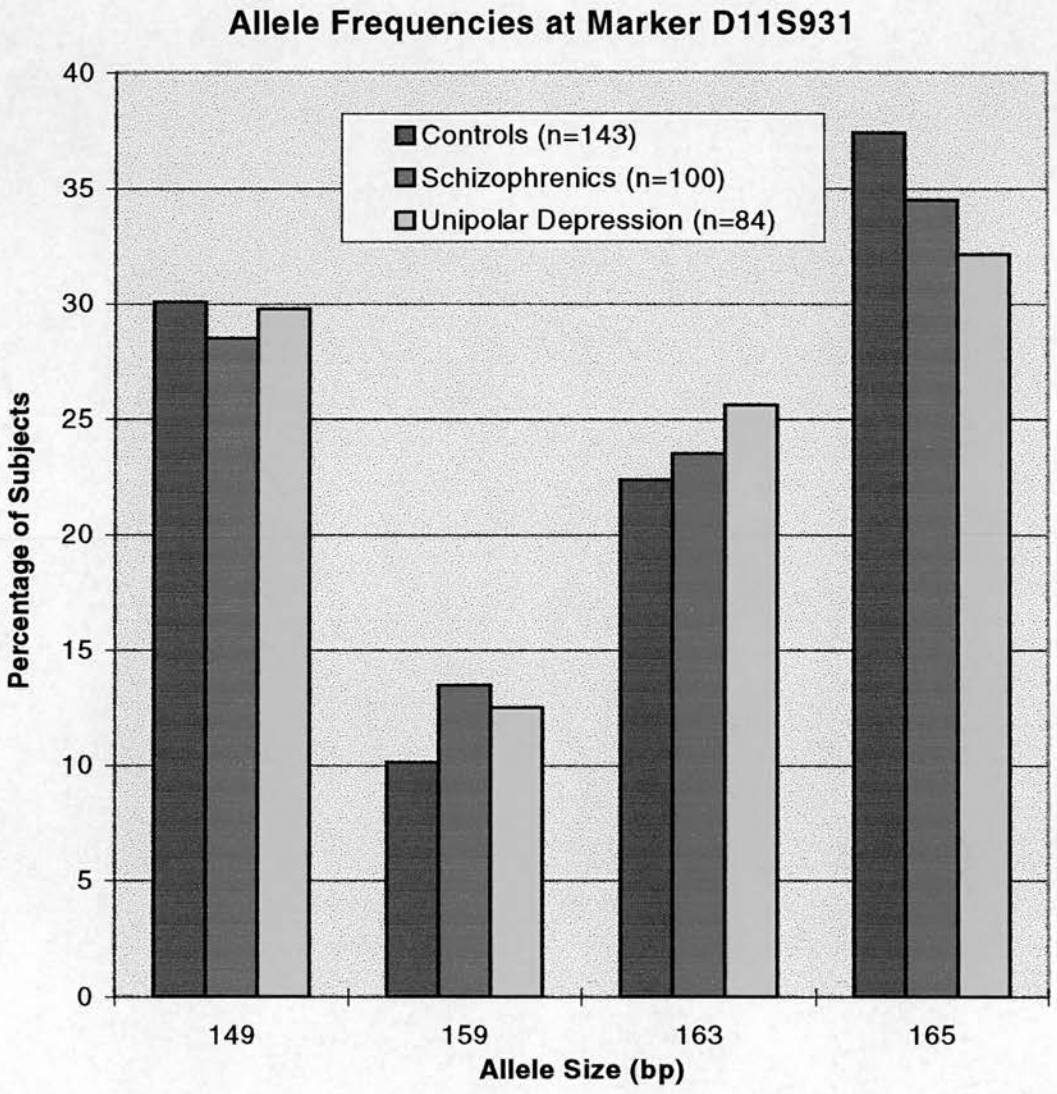


Figure 5.4 Graphical representation of percentage allele frequencies at marker D11S931 for the schizophrenia subjects, unipolar depressed subjects and control subjects.

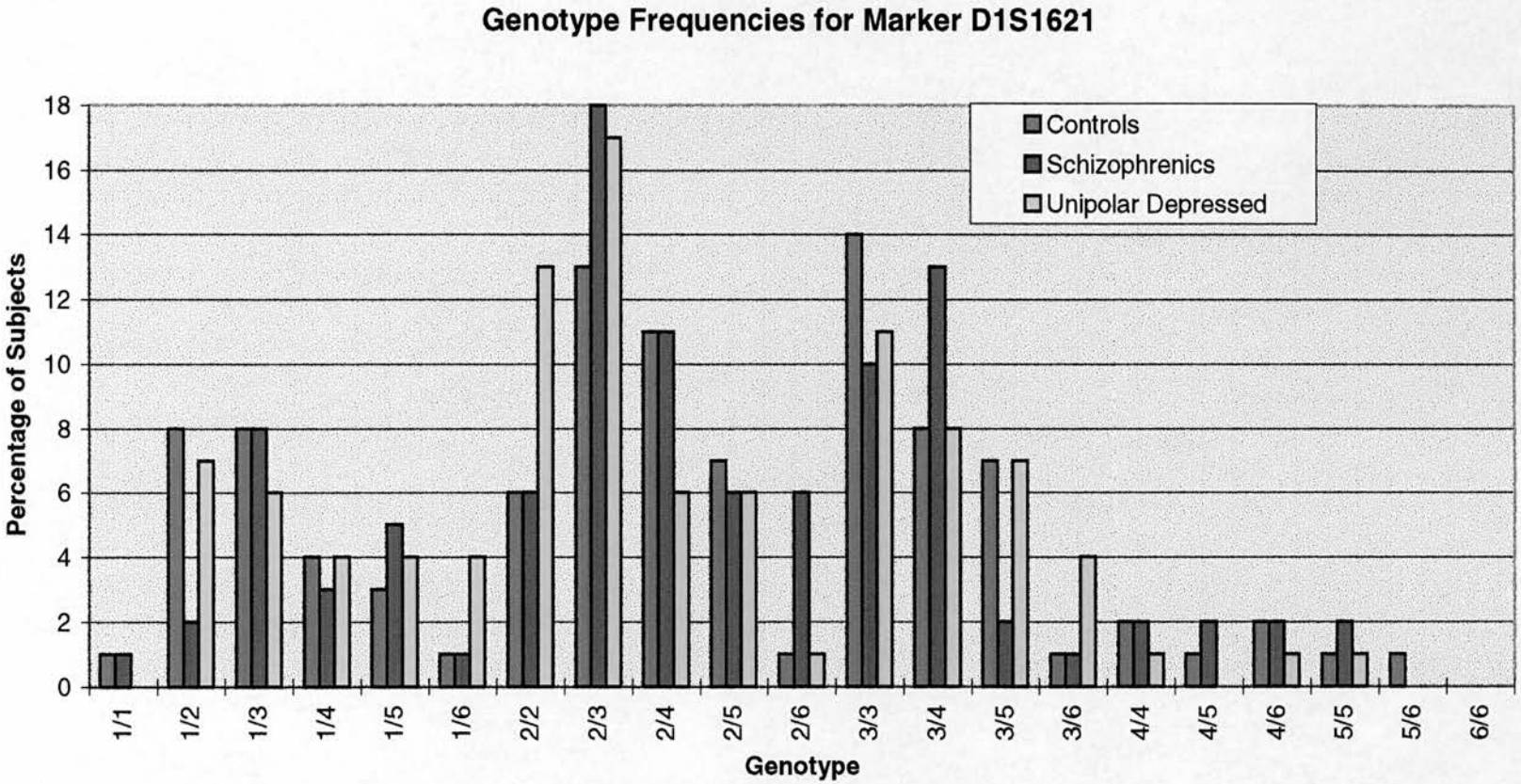


Figure 5.5: Graphical representation of percentage genotype frequencies at marker D1S1621 for schizophrenia subjects, unipolar depressed subjects and control subjects. Genotypes are based on the smallest allele being termed 1, therefore 1 corresponds to allele 252bp 2=255bp 3=258bp 4=261bp 5=264bp and 6=267bp.

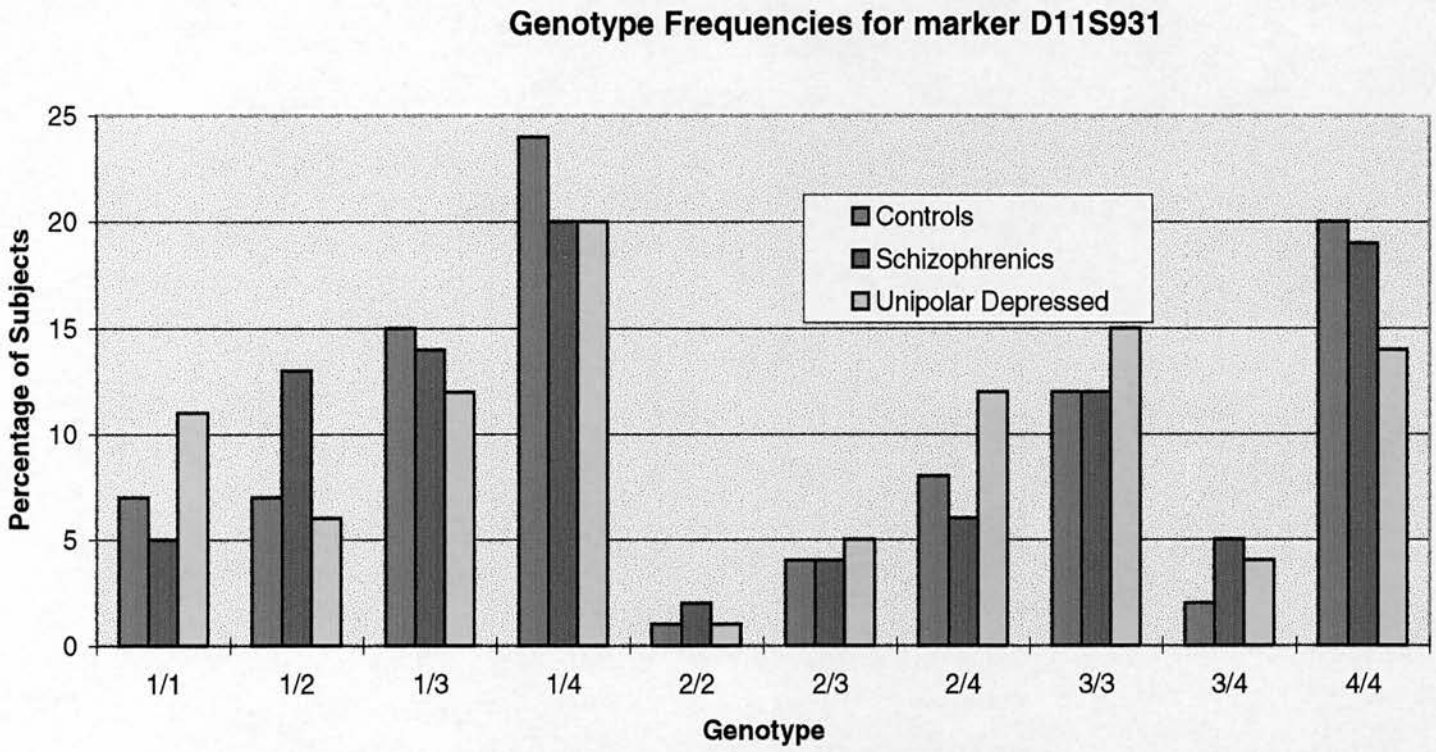


Figure 5.6: Graphical representation of percentage genotype frequencies at marker D11S931 for schizophrenia subjects, unipolar depressed subjects and control subjects. Genotypes are based on the smallest allele being termed 1, therefore 1= 149bp 2=159bp 3=163bp 4=165bp.

5.3.3 Statistical Analysis.

The presence of an allelic association was sought by using the computer program CLUMP (Sham and Curtis 1995).

This program is based on chi-squared analysis of the data. CLUMP is designed to assess the significance of the departure of observed values in a contingency table, from the expected values, conditional on the marginal totals. The program works on 2XN tables and was designed for use in genetic case-control association studies. The chi-squared value associated with the contingency table is defined in the usual way as the sum over all cells of the squared difference between the observed and expected value divided by the expected value. The significance is assessed using a Monte Carlo approach, by performing repeated simulations to generate tables having the same marginal totals as the one under consideration, and counting the number of times that a chi-squared value associated with the real table is achieved by the randomly simulated data (Sham and Curtis 1995). The chi-squared statistic supplied from the original 2XN table input into the program is termed T1.

CLUMP also provides a novel chi-squared value produced by "clumping" columns together in a new 2X2 table in a way designed to maximise the chi-squared value. The significance is again tested by the Monte Carlo simulations and the output chi-squared statistic is termed T4.

Statistical analysis on the association study data is shown in Tables 5.3 and 5.4.

Table 5.3: Statistical analysis of **allele** frequency data using the CLUMP computer package. T1 indicates normal chi-squared values and T4 indicates maximised chi-squared values after "Clumping" of data (see text). All values calculated from 1000 simulations.

	Chi-squared (χ^2)		P Value	
D1S1621	**T1**	**T4**	**T1**	**T4**
Schizophrenia subjects versus Control subjects	3.44	1.65	0.65	0.81
Unipolar Depressed subjects Versus Control subjects	5.29	3.44	0.39	0.46
D11S931				
Schizophrenia subjects versus Control subjects	1.59	1.30	0.66	0.69
Unipolar Depressed subjects Versus Control subjects	1.83	1.46	0.63	0.63

Table 5.4: Statistical analysis of **genotype** frequency data using the CLUMP computer package. T1 indicates normal chi-squared values and T4 indicates maximised chi-squared values after "Clumping" of data (see text). All values calculated from 1000 simulations.

	Chi-squared (χ^2)		P Value	
D1S1621	**T1**	**T4**	**T1**	**T4**
Schizophrenia subjects versus Control subjects	11.10	6.94	0.70	0.69
Unipolar Depressed subjects Versus Control subjects	10.57	6.80	0.75	0.71
D11S931				
Schizophrenia subjects versus Control subjects	5.78	5.10	0.77	0.56
Unipolar Depressed subjects Versus Control subjects	4.8	4.4	0.85	0.67

No significant statistical differences in allele or genotype frequency were found for markers D1S1621 or D11S931 in either the schizophrenia subjects or unipolar depressed subjects as compared to the control subjects.

The number of subjects used in this study was relatively small and some of the genotype groups were removed from the analysis or merged with other groups as their numbers were less than 5 thereby nullifying the chi-squared statistic. Further analysis of the data has been somewhat inhibited by the small numbers seen in the allele groups. Division of the data into sub-groups (such as age of onset or those subjects with a first degree relative with schizophrenia and those without) to look for an effect of association was not feasible due to the small number of subjects used in the study.

The D11S931 sample had 99% power to detect a 30% difference in allele frequency; a 76% power to a detect 20% difference and 23% power to detect a 10% difference.

The D1S1621 sample has 97% power to detect 30% difference in allele frequency, 68% power to a detect 20% difference and 19% power to detect a 10% difference (Cohen 1988).

Discussion:

Association studies which compare allelic frequency among unrelated affected and unaffected individuals from a population, are currently being widely used in the hunt for genes involved in psychiatric disorders.

There are essentially three possible explanations for demonstrating a significantly different distribution of allelic variants in affected versus unaffected individuals. The first of these is that the disease allele is in very tight linkage disequilibrium with the marker, the second is that the marker allele is actually involved in causing susceptibility to the disorder under study (pleiotropy) and the third and most disturbing is that the result is artefactual arising from population stratification. Population stratification occurs in a mixed population where any trait present at a higher frequency in a component of the population will show positive association with any allele which also happens to be more common in that group. To combat this problem homogeneous populations should be used where the control and affected group are closely matched for ethnicity. Alternatively, Falk and Rubinsteins (1987) Haplotype Relative Risk method in which the parental alleles that are not inherited by the proband, are used as the control sample provides an elegant solution to the problem of population stratification.

We observed no significant differences in the allele frequencies of markers D1S1621 or D11S931 in our population of schizophrenic subjects and unipolar depressed subjects in comparison to the control population.

This lack of association may indicate that the two markers investigated are not in linkage disequilibrium with a major gene of high prevalence in the population tested. This raises two issues the first of which relates to the linkage disequilibrium. The physical distance over which allelic association may be detected is dependent on several factors including recombination frequency and the number of generations from the founder. Data published using highly polymorphic markers suggests that associations may be detectable at distances of 1Mb or more (Peterson et al 1995). The hypothesis on which the current work is based is that a gene or gene(s) near the translocation breakpoint on chromosome 1 and/or 11 is disrupted in some way

by the translocation event and thereby causes the psychiatric diagnosis in the K26 pedigree. Several mechanisms may be responsible for the effect on the susceptibility gene. For example a gene which spans the translocation could be directly disrupted by the translocation event, making it non-functional or causing inappropriate splicing of two gene products resulting in a protein product which assumes a new function, or gene control elements may be removed or repositioned by the translocation event. Since the two markers used in the current study also lie in close proximity to the breakpoint, especially so in the case of marker D1S1621, it is likely that they will be in close proximity to the postulated schizophrenia susceptibility gene in accordance with our hypothesis. Thus, the chance of the postulated gene and the marker being in linkage disequilibrium and showing an allelic association is increased by choosing markers in close proximity to the breakpoint on chromosome 1 and 11. The negative result from the association study may reflect the lack of linkage disequilibrium between the markers and the postulated gene or indicate that the gene may not lie in close proximity to the breakpoint and is not in linkage disequilibrium with either of the markers.

Alternatively, it may be that this postulated gene in close proximity to the translocation breakpoint, is of major effect in the translocation family but is prominent in only a minority of the affected population studied and is therefore not of strong enough effect to be picked up in the association study. The lack of genetic linkage data published in the regions of the breakpoint perhaps reflects this hypothesis.

The number of subjects used in the current study has been relatively small and as a consequence the power of the association study to detect a gene of small effect will have been reduced. Ideally far larger numbers of subjects should be used in order to improve the statistical power of the technique to determine the presence of small gene effects. Given the absence of positive linkage on chromosome 1 it is likely that any gene in this region will have a small effect size. The sample size used in this study gives approximately 20% power to detect a 10% difference in allele frequency. Allelic association may be present but not have been detected due to too small a sample size.

It is unlikely that we have failed to demonstrate an association due to population stratification as the control and affected populations have been selected from the same geographical region as far as possible. Family based controls should however

be considered wherever possible to lessen this problem (Rubinstein *et al* 1981, Falk and Rubinstein 1987).

In conclusion, the present study does not reveal an association between the markers D1S1621 and D11S931 with either schizophrenia or unipolar depression in this Scottish population, evidence against a nearby gene of major effect.

Chapter Six

Isolation of genes around the chromosome 1 breakpoint by Coincident Sequence Cloning (CSC)

6 Introduction

From positional cloning on chromosome 1, it had been established that two cosmids (B01519 and I0142) and a PAC (DJ4B9/3) either spanned or resided in close proximity to the translocation breakpoint (work carried out by Kirsty Millar, Susan Anderson and Sheila Christie) (see Figure 6.1). In order to try to explain the significance of the translocation event in relation to the psychiatric illness in the K26 pedigree it was important to search for genes in this region. A direct cDNA selection technique, coincident sequence cloning (CSC)(Brookes et al 1994), was employed to achieve this aim.

Coincident sequence cloning is a technique which allows the selective isolation of DNA which is shared between two sources of DNA. If one of the sources of DNA is genomic DNA and the other cDNA then the coincident DNA between these two resources should be genes transcribed from the region of genomic DNA. There are two kinds of CSC. In the following experiments hybrid fishing (HF) and end ligation (EL) CSC have both been employed using a "one tube" procedure which is described below.

A schematic diagram of the coincident sequence cloning (CSC) method end ligation and hybrid fishing techniques and a one tube procedure combining these two techniques is shown in Figures, 6.2A and 6.2B and 6.2C.

The two types of CSC employed in this experiment, hybrid fishing CSC and End ligation CSC, differ in their stringency with the latter being a more stringent method due to end sequence specificity. Initially both methods are the same requiring two resources, one genomic and the other cDNA from the region which you wish to identify genes (e.g. cosmid DNA and foetal brain cDNA to identify brain expressed sequences from the cosmid). These two resources are cut with restriction enzymes and have catch linkers (synthetic oligonucleotide sequences represented in Figure 6.2 in pink and black) ligated to the ends to allow PCR amplification. In the case of the genomic resource, a biotin moiety (black circle) is added to allow physical isolation of this DNA using streptavidin coated magnetic beads at a later stage in the procedure.

In the hybrid fishing experiment (Figure 6.2A), the two resources are denatured and allowed to reanneal, forming homo- and heteroduplexes. Stringent washing stages remove the vast majority of non-coincident DNA before isolation of the DNA using streptavidin coated magnetic beads. An elution step allows the cDNA to be used as

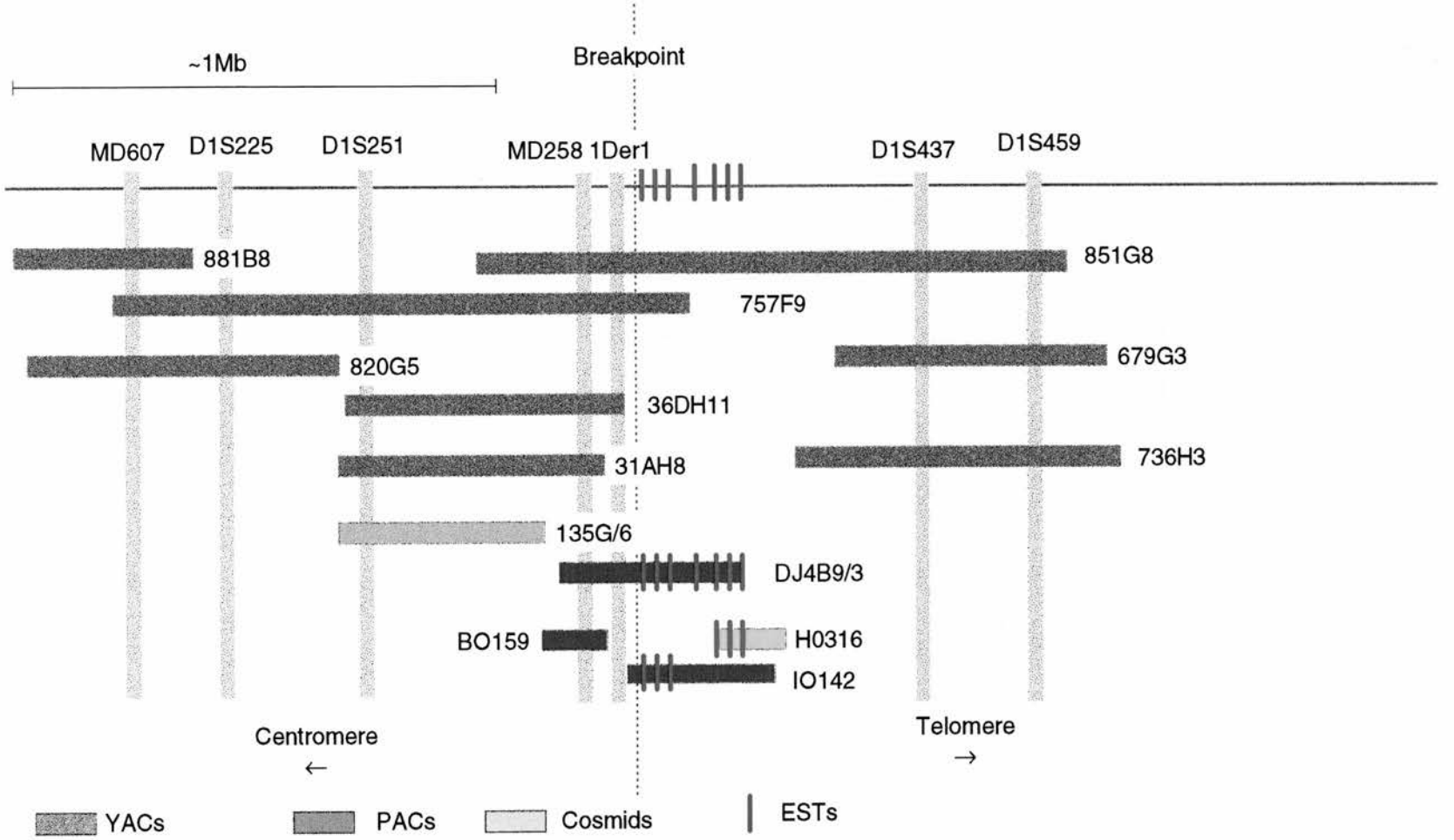


Figure 6.1; Chromosome 1 YAC contiguous clone map. Cosmids and PAC used in the CSC experiment are highlighted in blue.

template in a PCR reaction utilising primers complementary to the catch linkers. This is represented in the diagram as PCR 2.

In end ligation CSC, the catch linkers originally used to amplify the cDNA resource are removed by restriction enzyme digestion before heteroduplex formation. Once the complementary DNA sequences align, capture oligonucleotides (represented in Figure 6.2B and C as thin black lines (see key in diagram)), which are complementary in sequence to the initial catch linker sequence (thick black lines), used to amplify the genomic resource, are allowed to anneal. These capture oligonucleotides must be perfectly base paired and end matched before they can be ligated to cDNA sequences by a high temperature ligation step. This requirement for perfect base pairing and end matching makes end ligation CSC a more stringent procedure than hybrid fishing. The coincident DNA can be recovered as previously described and PCR reactions carried out after elution of the cDNA from the streptavidin coated magnetic beads, using primers which are complementary to the capture oligonucleotide sequences (PCR 1). Coincident products can then be cloned and analysed.

The technique which allows the combination of these two techniques in one experiment is represented in Figure 6.2Cand relies on the modified cDNA resource being present both with and without catch linker sequence (pink lines). The procedure followed is that of the end ligation technique but the hybrid fishing products can be amplified from the end product by primers complementary to the catch linker sequence. Separate product libraries can then be made from both the end ligation and hybrid fishing techniques and can be analysed.

The genomic resource in the current study was a pool consisting of cosmids I0142 and B0159 and PAC DJ4B9/3. Cosmid I0142 and PAC DJ4B9/3 had been identified as spanning the translocation breakpoint on chromosome 1 and B0159 had been shown to reside in close proximity to the breakpoint by DNA marker analysis. These two cosmids and PAC provide good coverage of the breakpoint region. The cDNA resource consisted of foetal brain cDNA which had been made from three ages of foetal brain tissue which were pooled (see section 2.15-2.17 for further details).

The CSC technique is capable of handling highly complex DNA resources and can enrich for coincident elements by up to 10^5 fold in the case of the hybrid fishing technique and 10^7 fold in the End Ligation technique (Brookes 1994). CSC is therefore a powerful method by which to identify transcribed sequences from a defined genomic region.

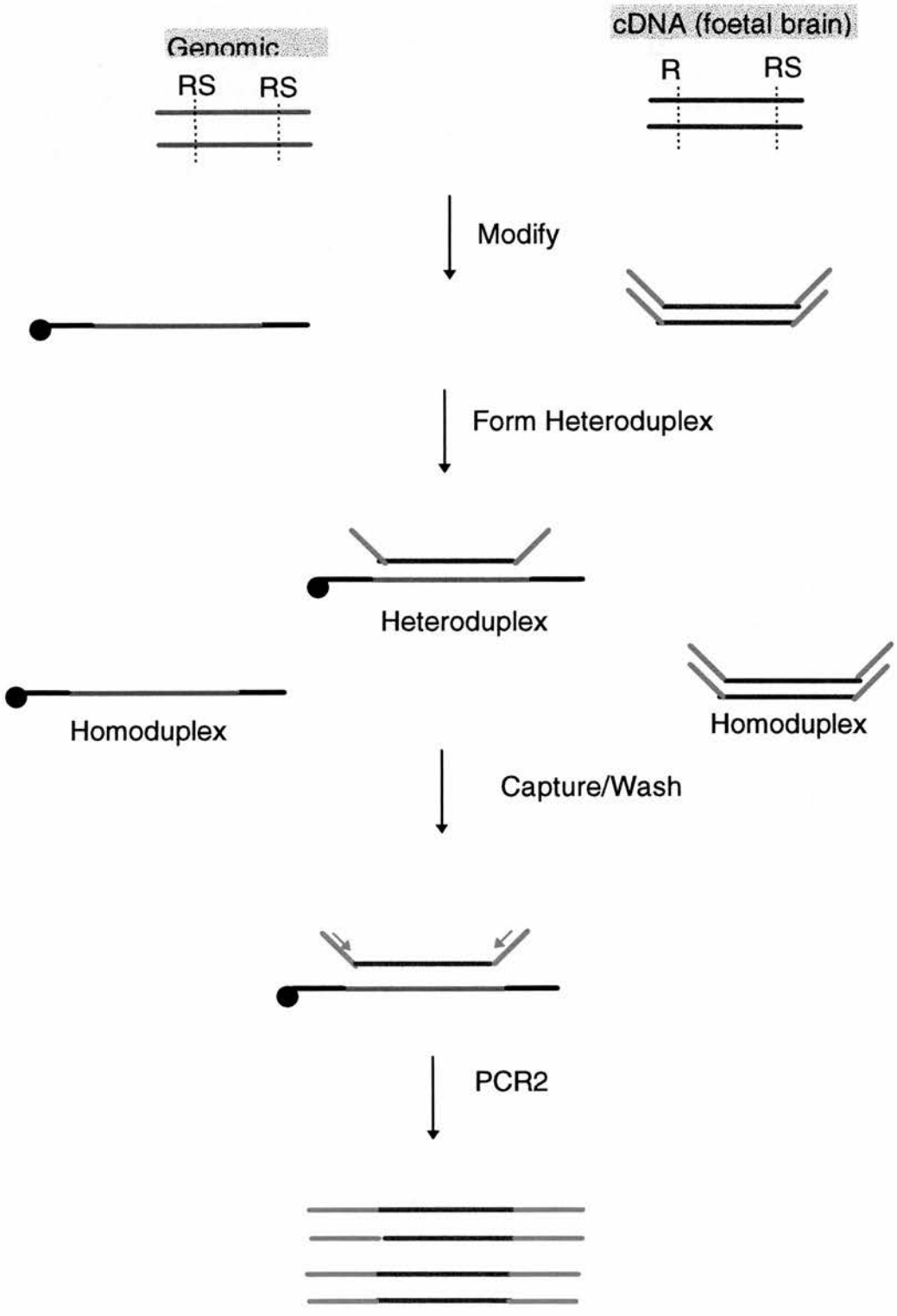


Figure 6.2A Schematic representation of Hybrid Fishing Coincident Sequence Cloning

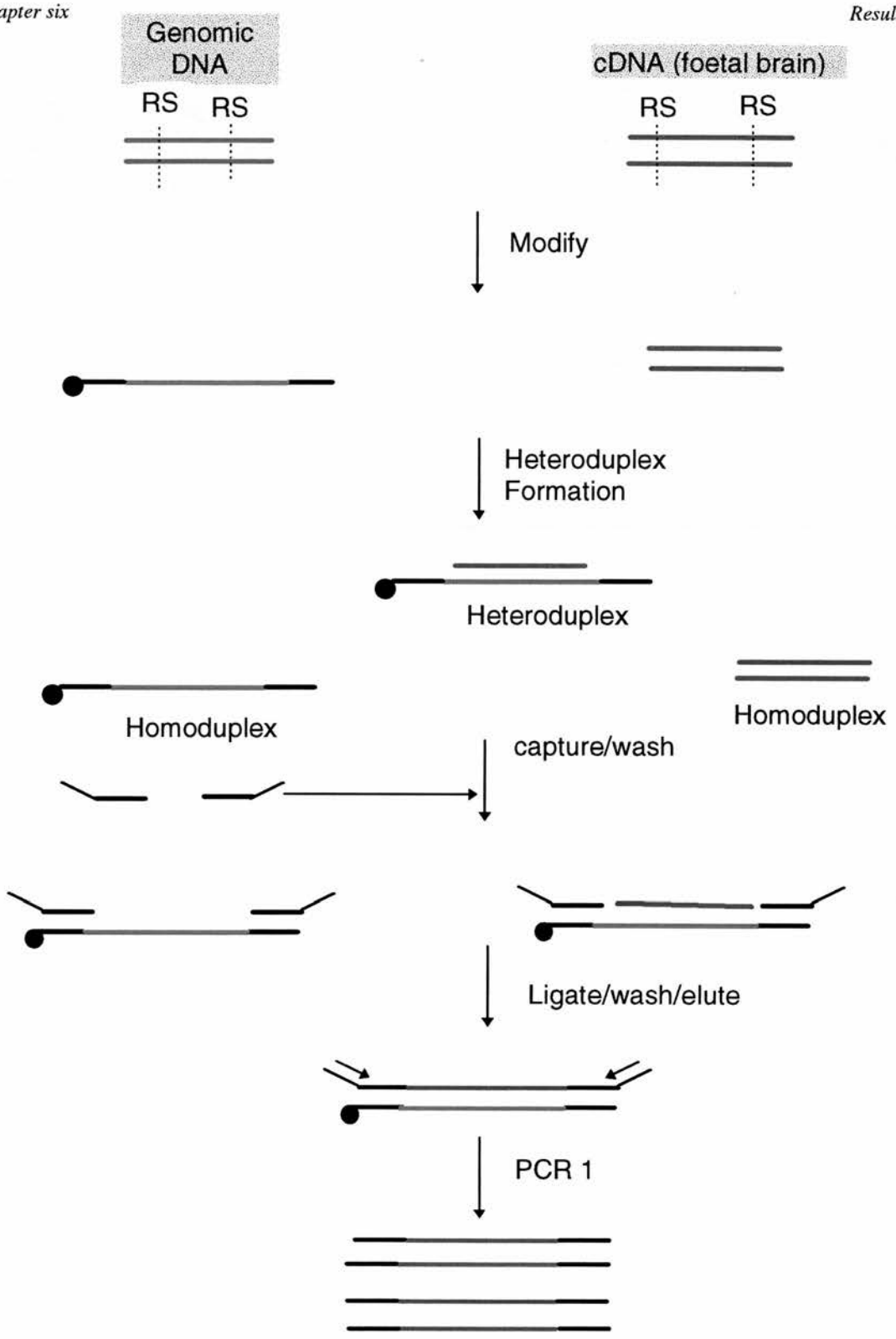


Figure 6.2B Schema of End Ligation Coincident Sequence Cloning

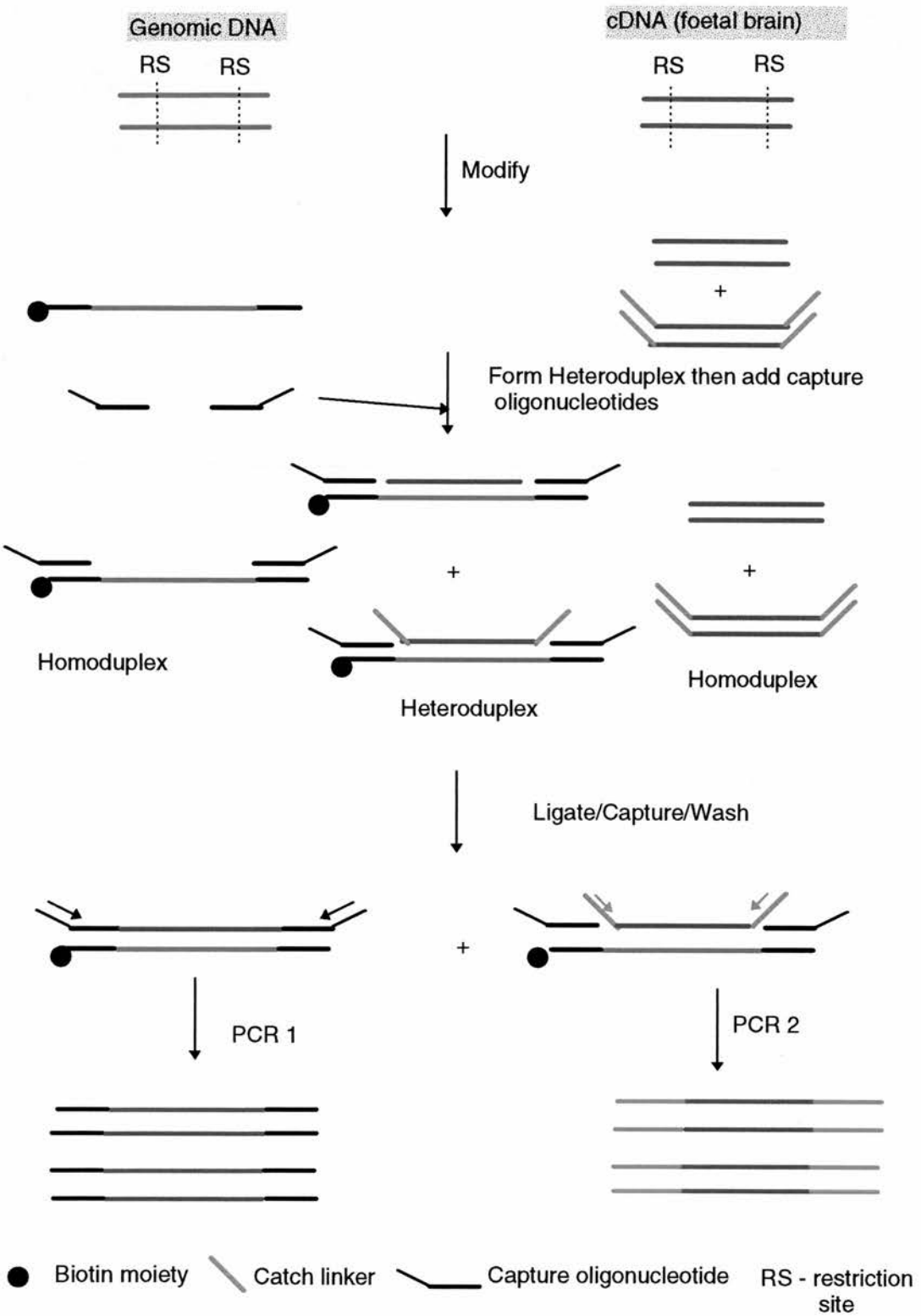


Figure 6.2C Schema of Coincident Sequence Cloning technique. End Ligation (EL) and Hybrid Fishing (HF) technique combined in one stage method. Left hand side represents End Ligation technique and right hand side represents hybrid fishing technique. See text and Figures 6.2A and B for further explanation.

6.1 Amplification of the product cDNA

The two input DNA resources were integrated as described in section 2.17.4 and the cDNA product eluted after stringent washing stages. This cDNA product was amplified using the conditions described in Table 6.1 using primer 727 (5'-GCGGAATTCTAGACTGCAGG -3') to obtain cDNA products from the hybrid fishing experiment and with primers 789 (5'- CGTTCCGGTAGCACGGG -3') and 596 (5'-GGACGGGTCGACACGCGAGG -3') and conditions described in Table 6.2 to obtain the End Ligation products. Two negative controls which had been through the streptavidin bead separation and washing stages were also amplified. The first of these negative controls was a "cDNA only" control in which a small aliquot of cDNA alone was put through the bead separation/washing stages to ensure that these stages were adequate and no unbound cDNA was left on the beads after the washing stages. The second control was a "no DNA" control, in which water was substituted for the DNA in order to determine if any contamination had occurred during the procedure.

Table 6.1: PCR conditions for amplification of **Hybrid Fishing** CSC cDNA products.

	Denaturation	Annealing	Extension
Cycle 1	94°C/30sec	54°C/30sec	72°C/1min30sec
Cycle 2-10	94°C/15sec	54°C/30sec	72°C/1min30sec
Cycle 11-20	94°C/15sec	54°C/30sec	72°C/2min30sec
Cycle 21-29	94°C/15sec	54°C/30sec	72°C/4min
Cycle 30	94°C/15sec	54°C/30sec	72°C/8min

Table 6.2: PCR conditions for amplification of **End Ligation** CSC cDNA products.

	Denaturation	Annealing	Extension
Cycle 1	94°C/30sec	55°C/30sec	72°C/1min30sec
Cycle 2-10	94°C/15sec	55°C/30sec	72°C/1min30sec
Cycle 11-20	94°C/15sec	55°C/30sec	72°C/2min30sec
Cycle 21-34	94°C/15sec	55°C/30sec	72°C/4min
Cycle 35	94°C/15sec	55°C/30sec	72°C/8min

PCR products were run out on a 2% agarose gel along with negative controls (Figure 6.3 and 6.4).

Hybrid Fishing Experiment:

Figure 6.3 shows the hybrid fishing cDNA PCR products obtained using primer 727. Although the "no DNA" control is negative, the "cDNA only" control has a strong smear in it similar to that seen in the positive integrated genomic and cDNA lane. Since the cDNA resource has no biotin moiety (as this is on the genomic resource only) it should not stick to the beads but should be washed off before the bound DNA is eluted. The presence of a smear in the negative "cDNA only" control indicated that either the washing stage was not adequate or that the cDNA was binding non specifically to the streptavidin coated magnetic beads. In order to try to combat this problem, the experiment was repeated using a more vigorous washing technique and also pre-blocking the beads with more sonicated salmon sperm and PVP (see 2.17.4) to try to ensure that all the non-specific sites on the beads were occupied. This however made little difference to the result, which still showed a smear in the cDNA only control (data not shown). The product in the positive integrated genomic/cDNA lane was not cloned due to difficulties in differentiating between the cDNA population which had come through from the beads due to inappropriate binding/washing and the truly coincident cDNA population from the experiment.

End Ligation Experiment:

Figure 6.4 shows the PCR results for the end ligation experiment using primers 596 and 789. The negative controls included in the PCR reaction are both negative and a smear is visible in the lane of integrated genomic and cDNA (refer to figure 6.4). Primers 596 and 789 are specific to the capture oligonucleotide sequence and therefore would not be expected to amplify the cDNA only negative control. The cDNA resource is catch linkered with 727/731. This means that only a very small amount of non specific cDNA should be present in the positive integrated genomic/cDNA PCR product and the majority of the cDNA amplified should be true end ligation products as it must have successfully ligated to capture oligonucleotides in order to be amplified in the PCR reaction. This product was therefore cloned into pBS (bluescript II SK) vector.

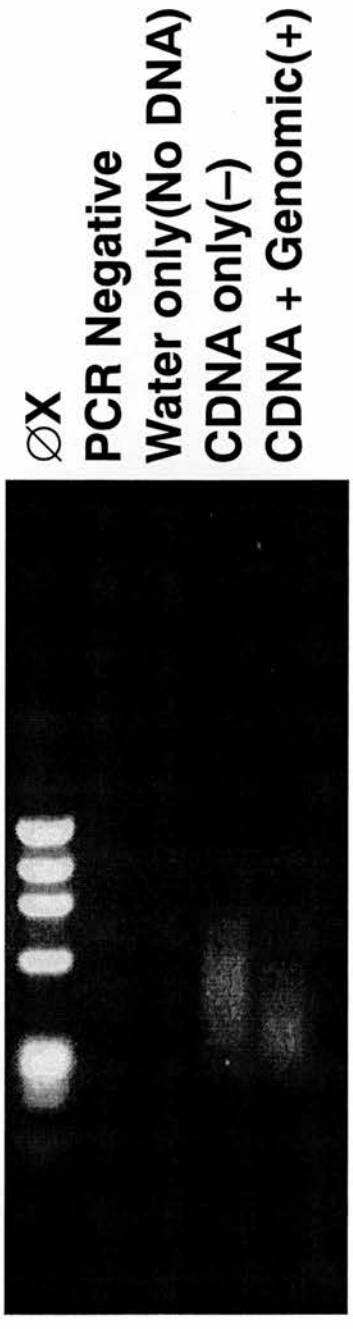


Figure 6.3: PCR products obtained from the **Hybrid Fishing CSC** experiment using primer 727. Negative controls - water only, cDNA only and a PCR negative control are shown as well as the integrated genomic and cDNA final PCR product.

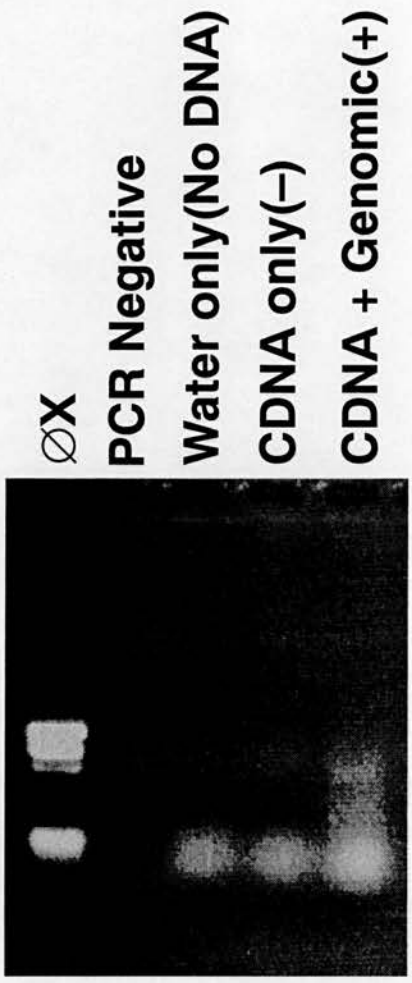


Figure 6.4: PCR product obtained from the **End Ligation CSC** experiment using primers 596 and 789. Negative controls - water only, cDNA only and a PCR negative control are shown as well as the integrated genomic and cDNA final PCR product.

6.2 Cloning of the End Ligation PCR product:

The capture oligonucleotide ligated to the cDNA in the CSC end ligation experiment incorporates an *EcoRI* site (GAATTC) which greatly facilitates the cloning step as cohesive end cloning can be employed. The amplified product was digested with *EcoRI* and cloned into pBS vector which had been digested with *EcoRI* and dephosphorylated. This vector was transformed into XL blue bacterial strain. The cells were plated at an optimum density so that individual colonies could be recognised and accurately picked by a colony picking robot (Hybaid). 1200 clones were picked and spotted onto five nitrocellulose filters in duplicate spotted format. Approximately 10% of the colonies contained no insert as determined by blue/white colour selection. Picking 1200 clones provided approximately two fold coverage of the region under study (~140kb).

6.3 Analysis of the End Ligation product library:

6.3.1 Strategy for analysis of End Ligation product library:

Initial analysis of the library included identification of clones containing high copy repetitive elements in order that these could be excluded from further analysis (section 6.3.2). Since there was no major band seen in the CSC product smear, 200 representative clones were picked and their inserts amplified by PCR in order to determine the average insert size of the clones. This was with a view to sequencing those clones with a reasonable insert size (over 100bp) by single pass automated cycle sequencing. The sequences from the ends of all clones were checked in order to ensure that the capture oligonucleotide and vector sequence were present. These clones were then analysed by computer database searches to see if they matched any previously identified products in the GENBANK and dbEST databases (HGMP resource centre, Hinxton Hall, Cambridge) or constituted novel cDNAs. Products with database matches to known human DNA sequences or novel clones were hybridised back onto the product library in order to determine how well they were represented in the library and to prevent repeated analysis of the same sequences. In order to prove that the clones were derived from the original input genomic resource, products were hybridised onto restriction enzyme digest panels of the cosmid and PAC genomic resource in Southern blot hybridisation analysis. To show that the clone was derived from the foetal brain cDNA either the clone was amplified from the original cDNA resource (which was very limited in supply) and/or

the clone was used to screen a full length foetal brain cDNA library (Clontech) which also served a secondary purpose of trying to extend the CSC clone sequence.

The CSC end ligation library was also screened with sequences which had been identified from positional cloning around the chromosome 1 breakpoint. A 7Kb wild type chromosome 1 fragment, which spans the breakpoint, had been identified by screening a genomic library made from a translocation patient DNA with a cosmid restriction fragment which had been shown to cross the breakpoint on chromosome 11. This hybridisation isolated a derived chromosome 1 fragment which was then used as a probe back onto the genomic library allowing the 7Kb wild-type chromosome 1 fragment to be obtained (see section 7.2). Another probe used to screen the CSC library was obtained by random sequencing and database searching of cloned restriction enzyme fragments (*Eag1 or Not1*) of the PAC DJ4B9/3 (work carried out by Kirsty Millar and Susan Anderson). One of the restriction enzyme clones identified by this procedure matched an EST in the database (GENBANK accession number N49833) and was used as a probe onto the CSC library. This clone resided within 30Kb of the breakpoint, as determined by restriction enzyme mapping, and resided outwith the 7Kb which the wild-type 1 sequence covered (Kirsty Millar and Susan Anderson).

The library was also screened with the PAC and cosmid inserts from the original genomic resources (see section 6.3.4).

6.3.2 Assessment of high copy repetitive elements in the library:

Repetitive elements are found in high frequency in genomic DNA (Alu elements comprise 5% of the human genome; Shen et al 1991) and in cDNA (3-5% of cDNA clones contain Alu elements; Crampton et al 1981). It is therefore likely that repetitive elements will be genuinely coincident products recovered from the experiment. Products which contain repetitive elements are problematic since the repetitive element can mask the effect of the non repetitive clone insert. The library was initially screened for repetitive elements by hybridisation analysis of gridded filters with Human Cot1 DNA, which is comprised almost entirely of high copy repetitive elements, primarily Alu elements. Prior to integration in the CSC experiment, the genomic and cDNA resources had been allowed to anneal to an excess of Cot1 DNA in order to quench most of the Alu elements to reduce the amount occurring in the product DNA.

The result of the Southern blot hybridisation analysis with human Cot1 DNA showed that approximately 2% of the library clones contained repetitive elements.

6.3.3 Analysis of End Ligation product library with probes residing in the vicinity of the chromosome 1 breakpoint:

The end ligation CSC product library was screened by hybridisation with the wild-type chromosome 1 7Kb fragment (which spans the chromosome 1 breakpoint) as well as with a DJ4B9/3 PAC restriction enzyme clone identified as an EST by sequencing and database searching (N49833)(see section 6.3.1). Results of the hybridisation analysis on the CSC library showed that for the 7Kb wild-type 1 fragment no positive clones were identified indicating that none of the products in the library resided within this 7Kb region around the breakpoint. Similarly for the PAC restriction enzyme clone, no positive results were obtained by hybridisation analysis suggesting that the EST matching clone (N49833) was also not represented within the CSC product library. (See discussion section for further explanation of these results).

6.3.4 Random sequencing of clones from the End Ligation CSC library

Boiled colony PCRs using primers designed to the pBS vector sequence (291; 5'CAGGAAACAGCTATGAC-3' and 292; 5'-GTAAAACGACGGCCAGT-3') were carried out on 200 clones from the End Ligation library in order to determine the average insert size and to identify clones with reasonably sized inserts (>100bp) for random sequencing. Table 6.3 lists the insert sizes found in these clones.

Table 6.3 Insert size of End Ligation CSC product library clones

Insert Size	Percentage of clones
<100bp	8%
100-200bp	73%
300-500bp	18%
>500bp	1%

Twenty clones with variously sized inserts were sequenced using primers 291 and 292. One of these clones (3A6) was very large ~2.3Kb. The sequence obtained was

checked against the GENBANK and dbEST databases and was found to have no database matches. This clone was therefore investigated further (see next section 6.3.5). One other clone had a very small insert composed of tandemly repeated capture oligonucleotide sequences only.

The remaining eighteen clones showed very strong identity to *E.coli* sequences (see Figure 6.5 for an example of the *E.coli* matches). These clones were however properly linkered with the capture oligonucleotide sequence being identifiable at one or both ends of the clone (it is possible that if the clone has an internal *EcoRI* site that the capture oligonucleotide sequence will be seen just at one site, the other end of the clone being ligated to the vector by virtue of the internal *EcoRI* site). This would indicate that the *E.coli* contamination has been present early on in the CSC procedure. The contamination has most probably originated from the PAC genomic DNA resource since it is unlikely that an artefactual product would be correctly linkered due to the stringency of the end ligation technique. PAC DNA is present in very low copy number in bacterial cells as the result of the vector origin of replication which only allows one copy of the PAC per cell. When the PAC DNA is purified the ratio of PAC DNA to *E.coli* background DNA is low in comparison with for example a plasmid DNA prep where there is a high copy number per cell. Consequently, problematic contamination from *E.coli* is more prominent when working with low copy PACs compared to higher copy plasmids or cosmids.

An explanation of how this *E.coli* contamination could have arisen from the genomic resource is given below.

The genomic resource is catch linkered in the initial stages of the CSC experiment with the following sequences; 477; 5'-CC**GAATTC**TAGAGTCGACC-3' and 479; 5'-GATCGGTCGACTCTA**GAATTC**GG-3' (bold lettering indicates the *EcoRI* restriction enzyme cutting site). The capture oligonucleotide sequences, which are ligated to the cDNA resource, are extended versions of these primers (5'-**GATCGGTCGACTCTAGAATTC**ACCCGTGCTACCGGAACG- 3' and 5'-GGACGGGTCGACACGCGAGGACC**GAATTCTAGAGTCGATC**-3' sequence in bold indicate the sequence identity to the catch linker sequences). Prior to ligation of the PCR amplified end ligation product (which has capture oligonucleotide sequences ligated to it) into pBS vector, the PCR product is cut with *EcoR*I restriction enzyme. Once this has been done the remaining sequence of the capture oligonucleotide and that of the catch linker sequence is identical. It is therefore not

possible to tell from this sequence being present in the final CSC product clone, whether this clone is a genuinely coincident product or is the result of contamination from the genomic resource. Southern blot hybridisation onto restriction enzyme digests of the cosmids PACs and original cDNA (or PCR analysis) is the only way to prove coincidence to both the DNA resources.

Figure 6.5A: CSC clone 100a sequence comparison with *E.coli* ada gene promoter region (M13155) as identified by GENBANK database searching.

```
100a:          78 GATCACCAAAGAAGGCGATAGCTTTAAAACCTGGATGTC 116
                  |||||||||||||||||||||||||||||||||||||||
E.coli:        29 GATCACCAAAGAAGGCGATAGCTTTAAAACCTGGATGTC 67


100a:      ACCACAGTTTAAAAGCTTCCTTGTCAGCGAAAAAAATTAAAGCGCAA 163
           ||||||||||||||||||||||||||||||||||||||||||||||
E.coli:    ACCACAGTTTAAAAGCTTCCTTGTCAGCGAAAAAAATTAAAGCGCAA 114


100a:      GATTGTTGGTTTTTGCGTGATGGTGACCGGGCAGCCTAAAGGCNATC 210
           |||||||||||||||||||||||||||||||||||||||||| |||
E.coli:    GATTGTTGGTTTTTGCGTGATGGTGACCGGGCAGCCTAAAGGCTATC 161


100a:      CCTAACCAGGGAGCTGATTATGAAAAAAGCCACATGCTTAACTGACG 257
           | ||||||||||||||||||||||||||||||||||||||||||||
E.coli     CTTAACCAGGGAGCTGATTATGAAAAAAGCCACATGCTTAACTGACG 208


100a:      ATCAACGCTGGCAATCTGTCTTAKCCCGCGACCCGAATGCCGACGGG 304
           |||||||||||||||||||||||+|||||||||||||||||||||
E.coli     ATCAACGCTGGCAATCTGTCTTAGCCCGCGACCCGAATGCCGACGGC 255

100a       GAATTCGATATC 316
           ||||||| | ||
E.coli     GAATTCGTTTTC 267
```

Figure 6.5B CSC Clone 15312 (247bp insert) sequence compared to the *E.coli* genome.

Escherichia coli K-12, complete genome
Length = 4,638,858

Vector sequence

```
15312 GTGGATCCCCCGGGCTGCAGGAATTCGCGGTCGGCATTCGGGTCGCGGGCTAA 121
      ||||  | | |||     | ||||||||||||||||||||||||||||||||
Ecoli GTGGTACGCACGGCGAAAACGAATTCGCCGTCGGCATTCGGGTCGCGGGCTAA 2308149

15312  GACAGATTGCCAGCGTTGATCGTCAGTTAAGCATGTGGCTTTTTTCATAA 171
       ||||||||||||||||||||||||||||||||||||||||||||||||||
Ecoli: GACAGATTGCCAGCGTTGATCGTCAGTTAAGCATGTGGCTTTTTTCATAA 2308199


15312   TCAGCTCCCTGGTTAAGGATAGCCTTTAGGCTGCCCGGTCACCATCACGC 221
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Ecoli:  TCAGCTCCCTGGTTAAGGATAGCCTTTAGGCTGCCCGGTCACCATCACGC 2308249


15312:  AAAAACCAACAATCTTGCGCTTTAATTTTTTTCGCTGACAAGGAAGCTTT 271
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Ecoli:  AAAAACCAACAATCTTGCGCTTTAATTTTTTTCGCTGACAAGGAAGCTTT 2308299

15312:  TAAACTGTGGTGACATCCAGGTTTTAAAGCTATCGCCTTCTTTGGTGATC 321
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Ecoli   TAAACTGTGGTGACATCCAGGTTTTAAAGCTATCGCCTTCTTTGGTGATC 2308349
```

Numbers on the left hand side indicate position in respective sequence at which this base resides.

If the *E.coli* contamination has been derived from the genomic resource then it must have been present in the initial stages of the CSC experiment. In order for the *E.coli* sequences, obtained from the CSC experiment, to have linker sequences present they must have been cut in the initial *Sau3AI* digest (to have compatible sticky ends onto which the linker sequence are ligated). The catch linker sequences would therefore have been added to both the genuine genomic DNA resource (i.e. the cosmids and PAC DNA) as well as to the *E.coli* DNA. This contaminating DNA would therefore be present throughout the CSC procedure. If these *E.coli* sequences constituted a large proportion of the "genomic resource," it is possible that when the genomic resource (with or without coincident product) was eluted off the streptavidin coated magnetic beads, only a small proportion of the DNA contained coincident cDNA. The PCR reaction would only amplify sequences which had the capture oligonucleotide ligated to them and not those with the catch linker sequences (since the PCR primers are designed to a unique segment of the capture oligonucleotide). When the end ligation PCR product was run out on a gel with the negative control samples (cDNA only and water only), a smear was only seen in the "integrated genomic and cDNA" lane of the gel and the "negative controls" were negative. However, a "genomic DNA only" control was not done so it is impossible to tell if a percentage of the smear in the positive lane is due to genomic only DNA which has come through the experiment. If this contamination was significant then it would also have been cloned and resulted in clones containing *E.coli* sequence inserts. It is therefore highly probable that the *E.coli* contamination has come from the PAC genomic resource.

From this preliminary sequencing data, it appeared that the end ligation CSC library was heavily contaminated with *E.coli* (90%; 18/20 clones) and was therefore going to be of limited use in identifying transcribed sequences. A further 30 clones were sequenced in order to see if the *E.coli* contamination was as prevalent as the initial data suggested and to see if any other novel human clones could be identified. Of these 30 clones only one (11B6) proved not to be *E.coli* by database searching, thus confirming the high level of contamination in the library. The clone 11B6 was further investigated (see section 6.3.5).

In order to try to find any truly coincident products amongst the plethora of *E.coli* containing clones the library was screened with the inserts from the B01519 and I0142 cosmids in order to try to identify any clones which were derived from these

resources. Since the cosmid vector (Lawrist 16) and the CSC clone vector (pBS) do not show any homology to each other (as demonstrated by sequence alignment analysis and database searches) the whole cosmids were used to probe the CSC library in Southern blot hybridisation analysis. This analysis resulted in a high background on the CSC library filters which may have been due to a small amount of *E.coli* contamination in the cosmid DNA preparation. Only one clone was identified as being positive in this hybridisation analysis. This clone (11G11) was sequenced and database searches showed that this clone insert matched lawrist 16 with 96% identity (see Figure 6.6). No linker sequences were present in the lawrist 16 sequence and it was probably ligated into the pBS CSC vector directly by virtue of the presence of *EcoRI* sites in its sequence. This sequence was therefore present at the *EcoRI* digestion step prior to ligation and represents an artefactual product and as such was not considered further.

6.4 Further analysis of the clones 3A6 and 11B6.

Clone 3A6:

Clone 3A6 has an insert size of 2.3Kb which is much larger than would be expected to result from a *Sau3AI* digested fragment (average size 300bp) and probably represents a partially digested *Sau3AI* clone. Sequence analysis of the ends of this clone insert, obtained using pBS vector primers 291/292, showed no matches to sequences in the GENBANK or dbEST databases.

- Is the CSC clone derived from the genomic resource?:

In parallel with the whole clone being sequenced, it was also important to show that the clone was truly coincident to both the starting resources. In order to demonstrate that the clone was derived from the genomic resource Southern blot analysis was carried out on restriction enzyme digests of the two cosmids I0142 and B01519 and the PAC DJ4B9/3 which constituted the original genomic resource. The whole cDNA clone was used as a probe and repetitive elements were blocked with sonicated total human genomic DNA. The result of this Southern blot analysis is displayed in Figure 6.7. Three restriction enzyme digests of the PAC and cosmids are present on the filter (*BamHI, Hind III* and *EcoRI*). In each digest of the PAC a band is present, but no bands were visible in either of the two cosmids, indicating that this

Figure 6.6: Sequence comparison of CSC clone 11G11, isolated by Southern blot hybridisation with cosmid B01519 insert, with Lawrist 16 vector sequence from GENBANK database.

```
11G11:                68   GAATTCTGGCGAATCCTCTGACCAGCCANA 98
                           |||||||||||||||||||||||||||| |
Lawrist:           46995   GAATTCTGGCGAATCCTCTGACCAGCCAGA 47025

11G11:       AAACGACCTTTCTGTGGTGAAACCGGATGCTGCAATTCAGAGCGG 143
             |||||||||||||||||||||||||||||||||||||||||||||
Lawrist:     AAACGACCTTTCTGTGGTGAAACCGGATGCTGCAATTCAGAGCGG 47070

11G11:       CAGCAAGTGGGGGACAGCATAANACCTGACCGCCGCAGAGTGGAT 188
             |||||||||||||||||| || ||||||||||||||||||||||
Lawrist:     CAGCAAGTGGGGGACAGCAGAAGACCTGACCGCCGCAGAGTGGAT 47115

11G11:        GTTTGACATGGTGAARACTATCGCACCATCAKCCAKAAAACCGAA 233
             |||||||||||||||+||||||||||||||+|||+|||||||||
Lawrist:     GTTTGACATGGTGAAGACTATCGCACCATCAGCCAGAAAACCGAA 47160

11G11:       TTTTGCTGGGTGGGCTAACGATATCCGCCTGATGCGTGAACGTGA 278
             ||||||||||||||||||||||||||||||||||||||||||||
Lawrist      TTTTGCTGGGTGGGCTAACGATATCCGCCTGATGCGTGAACGTGA 47205

11G11:       CGGACGTAACCACCGCGACATGTGTGTGCTGTTCCGCTGGGCATGC 323
             ||||||||||||||||||||||||||||||||||||||||||||
Lawrist:     CGGACGTAACCACCGCGACATGTGTGTGCTGTTCCGCTGGGCATGC 47251

11G11:       CAGGAAAACTTCTGGTCCGGTAACGTGCTGAGCCCSGCCAAACTCC 369
             ||||| ||||||||||||||||||||||||||||+||||||||||
Lawrist:     CAGGACAACTTCTGGTCCGGTAACGTGCTGAGCCCGGCCAAACTCC 47297

11G11:        GCGATAAGTGGACCCAACTCCAAATCAACCGTTAACAAKCAACAGG 415
             |||||||||||||||||| ||||  | |  ||||||+|||||||
Lawrist:     GCGATAAGTGGACCCAACTCGAAATNCAACCGTAACAAGCAACAGG 47342
```

```
K= T or G
M= A or C
R= G or A
S= C or G
W= A or T
Y= T or C
```

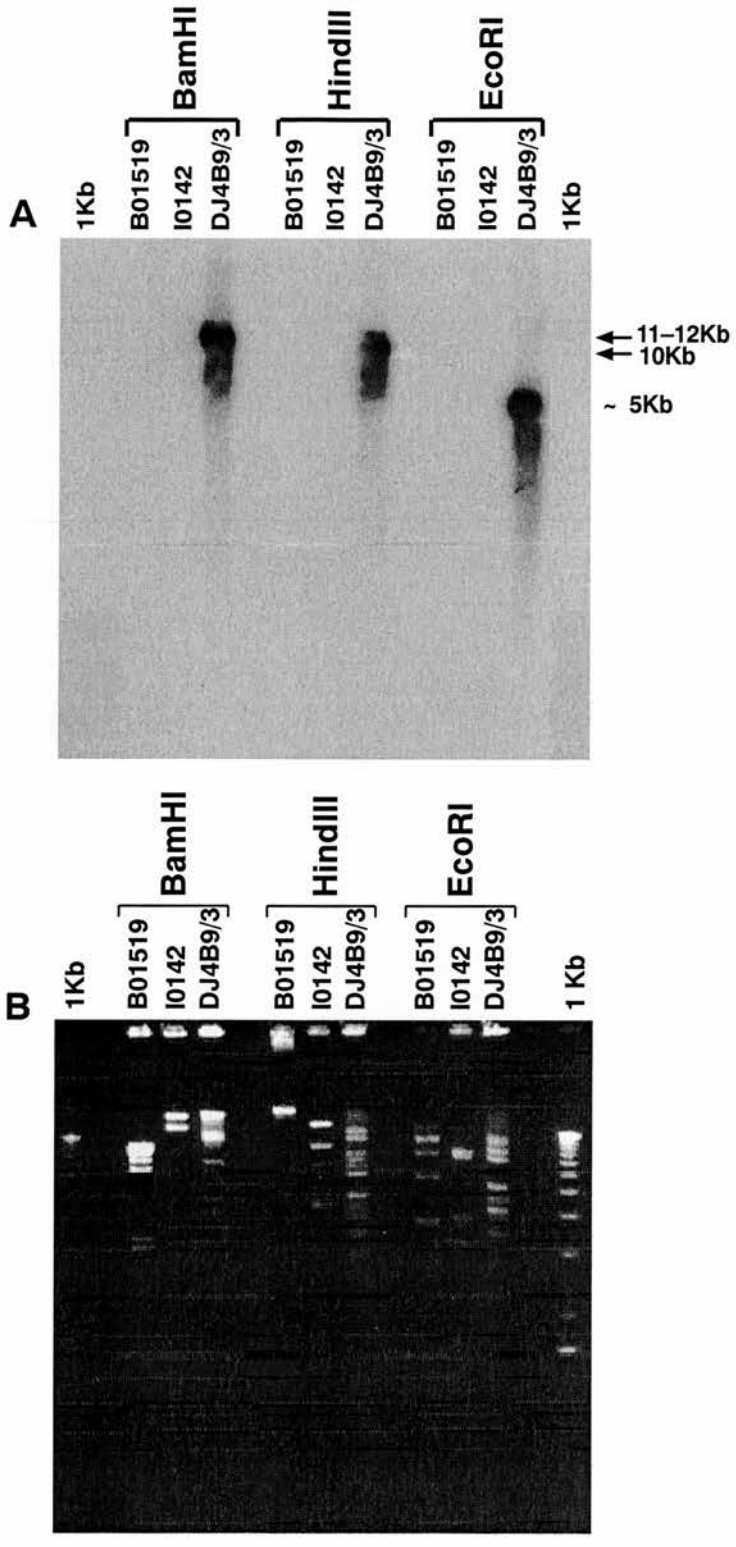


Figure 6.7: Southern blot hybridisation analysis of CSC clone 3A6 (whole clone insert) hybridised onto restriction enzyme digests of the cosmids BO1519 and I0142 and PAC DJ4B9/3.

clone is contained within the PAC only. This would place clone 3A6 distal to the chromosome 1 breakpoint. The 3A6 clone must also lie outwith the immediate vicinity of the breakpoint since it is not contained in the I0142 cosmid (refer to Figure 6.1). This cosmid spans the breakpoint and overlaps with PAC DJ4B9/3 in the region immediately proximal and distal to the breakpoint. If the clone mapped in this region it would therefore be expected to be seen in cosmid I0142 as well as the PAC. Similarly, the clone cannot map proximal to the breakpoint as the PAC overlaps with either cosmid I0142 and/or B01519 in this region. The bands seen in the PAC digests are contained within a background smear. It had previously been established, by hybridisation with Cot1 human cDNA, that this CSC clone contained repetitive elements. The background smear is present as a result of these repetitive elements not being completely quenched by the blocking procedure with sonicated total human genomic DNA.

- Is the CSC clone derived from original cDNA starting resource?:

Once it had been established that the clone had been derived from the original genomic resource it was also important to show that it had also come from the cDNA resource. In order to establish this, a PCR reaction was carried out on the original pooled cDNA resource using several primer pairs extending across the clone sequence. These primers were used in the sequencing of the whole CSC clone (see next section) and are described in Table 6.5 (M306/I184 → 598bp product, M188/M187 → 1865bp product, M305/M187 → 1501bp product and I184/M187 → 946bp product. See Figure 6.10 for the 3A6 clone sequence and the position of these primers). A positive and negative control were also included in the PCR reactions the positive control being the 3A6 clone DNA and the negative control being a water only control to show any contamination present. The PCR reaction using these primer sets failed to amplify a single band from the original starting cDNA although a band was seen in the 3A6 clone DNA positive control (Figure 6.8). This may indicate that the CSC clone was not derived from the initial cDNA and represents a genomic artefact or alternatively it may be that this clone is very low copy in the original cDNA and has not amplified well enough to be visible in the PCR product. A secondary PCR reaction was carried out using 1µl of the primary PCR product as template in the reaction. This PCR similarly failed to demonstrate a band of the correct size (not shown).

A foetal brain cDNA library (Clontech) was also screened with a probe produced by PCR with primers M306 and I184 on 3A6 clone DNA. The primary screening of this library was done at high stringency (68°C) and several positives clones were identified. These positives were weak, taking several days of exposure to film to be visible. These clones were purified to single plaques which were then subjected to boiled plaque PCR using primers designed to the vector sequences (391: 5'-AGCAAGTTCAGCCTGGTTAAGT-3' and 392: 5'- TTATGAGTATTTCTTCCAGGG - 3'). The conditions for the PCR are described in Table 6.4. The PCR products were then sequenced and compared to the sequence of the clone 3A6 to see if these cDNA clones further extended the sequence of the CSC clone. This procedure served two purposes, the first of which was to try to demonstrate that this DNA fragment corresponded to a gene fragment expressed in foetal brain tissue from which the original starting cDNA resource was derived. Secondly, the screening of the full length cDNA library provided an opportunity to further extend the sequence of the CSC clone. Extending this sequence would allow further information about the structure of this gene to be gleaned and to permit further database searches.

Table 6.4: PCR conditions for the amplification of cDNA clones from the foetal brain cDNA library.

	Denaturation	Annealing	Extension
Cycles 1-20	94°C/30secs	56°C/45secs	72°C/2mins
Cycles20-30	94°C/30secs	56°C/45secs	72°C/3mins

The sequences of a cDNA clones from the Clontech foetal brain cDNA library can be seen in Figure 6.9. These sequences did not extended the 3A6 clone sequence. When the sequences of the two clones, were compared to the consensus sequence for the 3A6 clone only between 60 and 70% identity was observed. The probe which was used to screen the foetal brain library was 598bp in length and contained 200bp of MALR repeat sequence (see Figure 6.10). When the sequence of the cDNA clones from the library screen were studied, they were shown to contain large repeat elements as demonstrated by "repeat masker" (Jurka et al 1992) repeat database. Homology between the repeat elements in the probe and in the cDNA clone were probably responsible for the clone being identified as positive. Since the match between the probe and cDNA clone was not

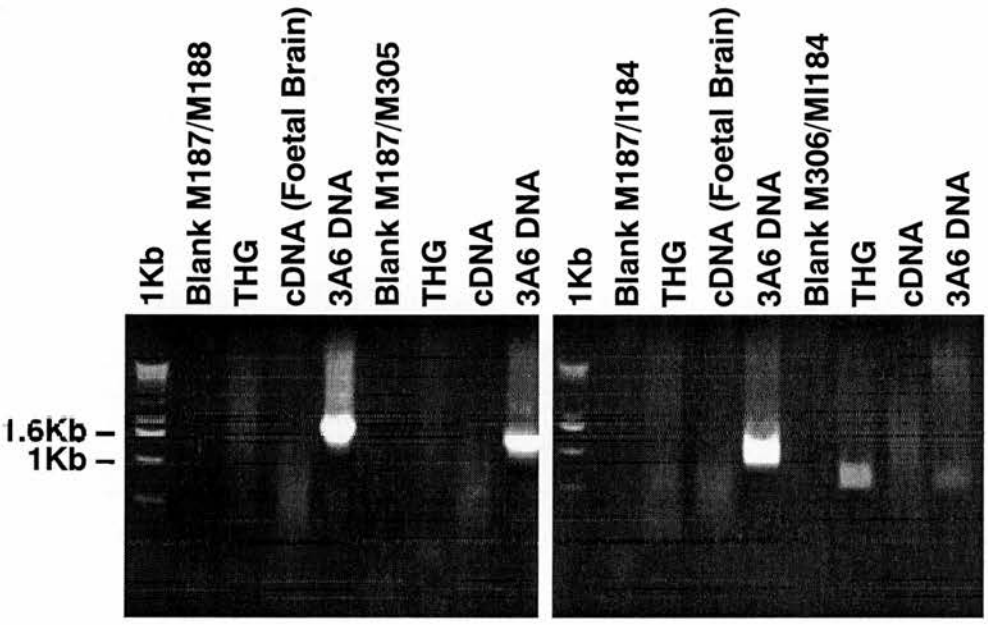


Figure 6.8: PCR analysis showing amplification of CSC clone 3A6 from original cDNA resource, CSC clone DNA and Total humanGenomic DNA(THG)

Figure 6.9: Sequence of cDNA clone 721obtained by screening a foetal brain cDNA library with CSC clone 3A6.

A) Sequence of cDNA clone 721

```
   1  TCCGGTTTTT  TTTTCACTCA  CCCCTCAATC  CAGAGAGGGC  ATTAATTTAT
  51  TCATGAGGGA  TCTTCCACCA  TGACCCAAAC  ACCTTCCATT  AGGCCCCACC
 101  TCCAACACTG  AGATCAAACT  TCAACATGAG  ATTTGGAGGG  ACAAATATCC
 151  AAACTATAGC  AACATATCAC  TTTTGTACCA  TCATAAAGCT  GAGAAATTGT
 201  TAGTTGACAT  TGTTGGGGAC  CATCTCTATA  ACAGTTCATT  GGAACAAGTC
 251  TGTTAAAAAT  CATGTTCATA  CTGATACACA  AAACATGTAA  TTATCATACA
 301  TCTAGTATCA  TGTCACACTC  AAAAGATAAA  TGGTTATAAA  TTTTACAAAA
 351  GGAGGCTGCA  GTTAGGAATG  TTTAATTTAC  AGTTCTTTCT  CACCTACTAT
 401  CAACAAACAG  TACCGGAATT  CArCTTGGrC  TTAACCAGGC  TGAACTTGCT
 451  TCTNNCNGTC  TG
```

MALR REPEAT

MER REPEAT

r = A or G

N= not determined

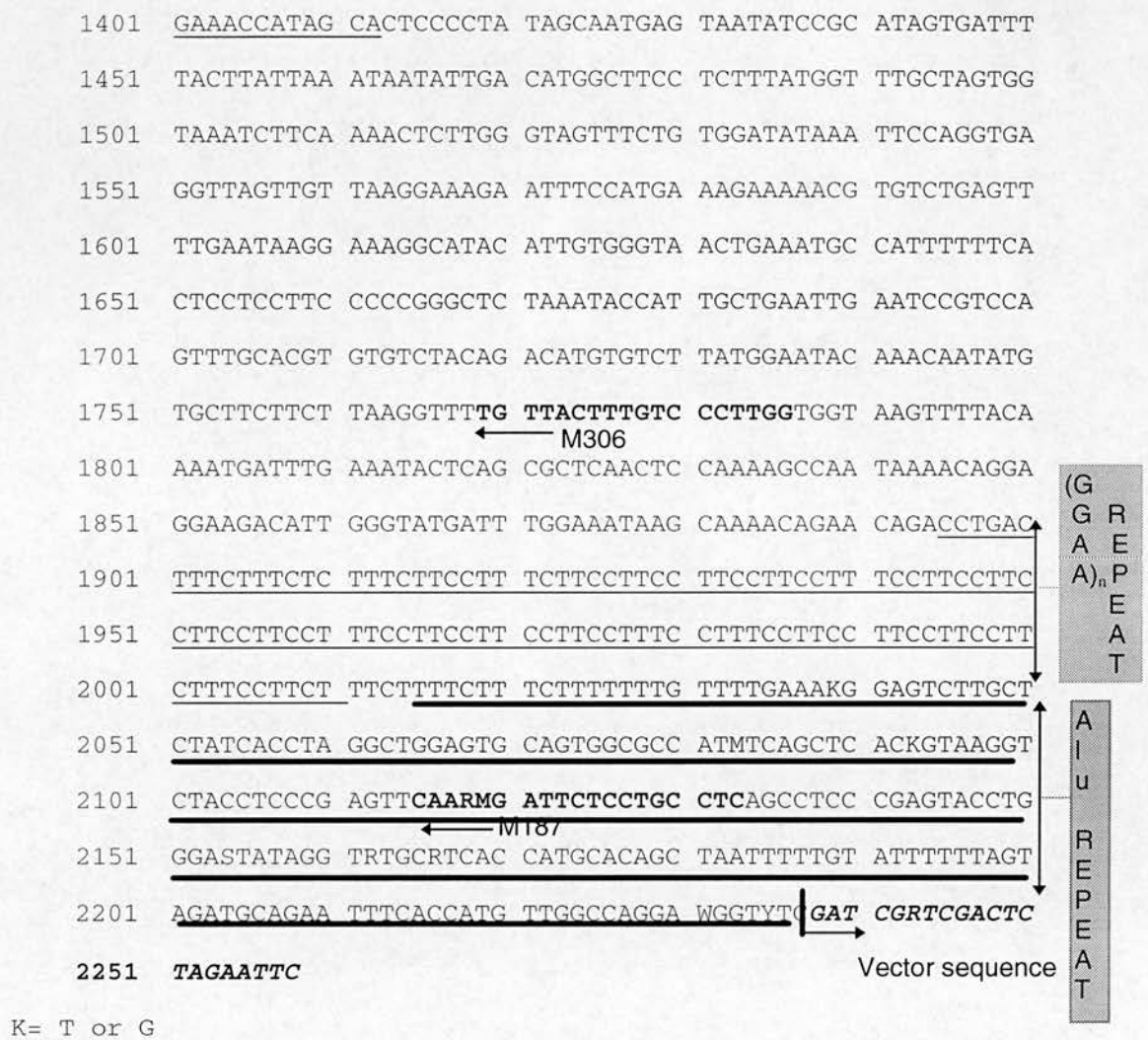
B) Sequence comparison of CSC clone 3A6 and cDNA clone 721. Sequence identity was 67.1% identity in 170 bp overlap.

```
721                                  TCCGGTTTTTTTTTCACTCACCCCTCAATCCAGA
                                       ||  |||| | |   |||||       |||
3A6     AAGAGAAAGCAAGCAAGGGGGGGCCAAGTTGTCCTTTTCTAATGAACCCCCTCCCACAGT
        1220      1230      1240      1250      1260      1270

            40        50        60        70        80        90
721     GAGGGCATTAATTTATTCATGAGGGATCTTCCACCATGACCCAAACACCTTCCATTAGGC
         | ||| |||||  ||||  |||||   | || ||||||||||||||| | |||||||||
3A6     AACGGCGTTAATCCATTCTAGAGGGTGGTGCCCCCATGACCCAAACACTTCCCATTAGGC
        1280      1290      1300      1310      1320      1330

           100                  110       120          130       140
721     CCCACCTCC-AACACT---------GAGATCAAACTTC---AACATGAGATTTGGAGGGA
        ||||||||| ||||||         | ||||| | |||    |||| |  ||||  ||||
3A6     CCCACCTCCTAACACTGCCACATTGGGGATCAGATTTCCAACACATTAACTTTG--GGGA
        1340      1350      1360      1370      1380      1390

              150       160       170       180       190       200
721     CAAATATCCAAACTATAGCAACATATCACTTTTGTACCATCATAAAGCTGAGAAATTGTT
        |  | ||  |||| ||||||
3A6     C--ACATTGAAACCATAGCACTCCCCTATAGCAATGAGTAATATCCGCATAGTGATTTTA
            1400      1410      1420      1430      1440      1450
```

Figure 6.10: Sequence of CSC clone 3A6. Capture oligonucleotide sequences are marked in bold italics at the beginning and end of the sequence and sequencing oligonucleotides are also in bold. Repeat elements are also marked on the sequence.

```
   1  GAATTCTAGA GTCGACCGAT CCGGTTTTAC AGGAGGAATT AACCTGGGAT
  51  GGGGGAAAGA ATGCATTTAG TGCTGGAATA TATAAGGAGA GGCTTTCCTT
 101  TGTCCCCTCC TCTCTTCTCT CTCCCCCAGG ATTTTTTGTC CTTCCGTGTT
 151  CCACCTGATC CATGGCCAAC TGGCTTTAGA AGTTGTTATT AATAGTTCAC
 201  CATGCTCATT TTCATGAAAC AGAGGGAAAG GCAGCCTCTT CATTTGATGT
 251  CTAGTTAGCA AACCAGARGT TTGTCACCAG CTGGTCATAA CAGACATTTT
                          →  M188
 301  TAATAAAACT TACCTCCMAG TTAAAATAAA AATGCTGCCT TAGAGATTTT
 351  GTCCTGAAAC ATACGGAAAA ATGACAGGTG GGATCTCTCA TGCAAAACTT
 401  CTGCTAGGGC TAAAGGAAAA TATCTGATAA GGTTGTTTGC CCATGGGGGA
 451  AGGATTAATT GGATTTTCTG TCTTGCCCCT AGTGTCTGAC CCTCTGTTAT
 501  TATGACAGGA CCACAGWAGT TTTGAACGCT GAATCATAGT TTATAGTTTT
 551  AGCATTTTTG AGCWATGGAG ATAAGAGTGA GGGAAGGAGT GTGATAGGGG
 601  AGTGAGTTCC CACAATACTC CGGTGATAGT TATTGTCTGC CTGAAAATGT
                                      M305  →
 651  GAATACATAT ATTCAAGCCC GATGTAGCCT GTTCGAAGCA ACCAGCCGGG
 701  CAATGAAAGC CTCAGCTGAA GATCATAGGA TTAAATCCAG GCTCCCCTCT
 751  TTCAAGTACC TTCTCATGCT ACTGCATTTA TGTGGGTAGT CTGGGATTAT
 801  GTCTGTGGTT TGTTTCACTA TTTTCTGGAT CACTTCTGAA CCAGTCCGTG
 851  TGTCAGTCCT GTGGATATGC ATTTCATGAA TGAGACTCTG AGGAATTTTC
 901  TAATAGTGTC TATGGTGGCA GTGGGAAAAG GGGAGGAGGC AGAGAAAGAG
 951  GTGGTAGCAA GAGAAAGAAC TGTGTGGAGC TCGAGCTAAA AGCAGGAGCT
1001  AAGACTGCAA GATCAAATGA TTCCCACTCA GTTGTCCTAG TCTCYGTCYT
1051  TGTTTTCTGT GGCCATAACA GAACACCATA GATTTGGTAA TTTATAAAGA
1101  GCAGAAACAT ATTTCCTCAT CATTCTGGAG ACTGGGAAGT TCAAGATCAA
1151  GGGCTGGCTC AACAGAAGCC TCTTGCTGCA TCATGCCATG GCAGAAGGTG
1201  GAAAGGSAGA GAGGGCAAGA GAAAGCAAGC AAGGGGGGGC CAAGTTGTCC
          →  M184
1251  TTTTCTAATG AACCCCCTCC CACAGTAACG GCGTTAATCC ATTCTAGAGG
1301  GTGGTGCCCC CATGACCCAA ACACTTCCCA TTAGGCCCCA CCTCCTAACA
1351  CTGCCACATT GGGGATCAGA TTTCCAACAC ATTAACTTTG GGGACACATT
```

LTR MALR REPEAT



K= T or G
M= A or C
R= G or A
S= C or G
W= A or T
Y= T or C

exact the positives seen on the autoradiogram were weak. Therefore, it appears that no real positive clones were detected from the foetal brain cDNA library.

As no true positive cDNAs were obtained from this library screening it is possible that the CSC clone is a genomic clone and not a cDNA clone. In order to further determine if this CSC clone was a genomic clone, a PCR was carried out which utilised the primers used to sequence the whole CSC clone. These primers were designed to amplify consecutively larger fragments of the clone (see Figure 6.10). By carrying out this PCR on CSC clone DNA as well as on total human genomic DNA and A9 (hybrid containing only human chromosome 1) DNA, it was possible to test for regions where the clone DNA, but not the genomic DNA, would be amplified. Such a result may be indicative of the presence of large intronic sequences across which the PCR cannot extend in the genomic DNA. Results from this PCR reaction are shown in Figure 6.11. The region of sequence which constituted the probe used to screen the cDNA library was amplified from both the genomic resource and from the clone DNA (M306/M184). However, as the fragments amplified extended further across the clone sequence a product was still visible in the CSC clone DNA (3A6), but not in the genomic DNA. This may be the reflection of large intronic sequences existing between the primers in the genomic, but not the CSC clone DNA, or may simply reflect the higher complexity of the genomic DNA, compared to that of the CSC clone DNA. Sequencing the genomic DNA directly to look for points of divergence and splice sites would distinguish these possibilities.

It is also possible, since this CSC clone is so large and contains several internal *Sau3AI* sites, that it is chimaeric, i.e. made up of several different unrelated *Sau3AI* fragments which have been spuriously ligated together. This could be another reason that the genomic PCR did not work as the primer pairs extended further across the clone sequence and may explain why this clone could not be amplified from the original cDNA resource.

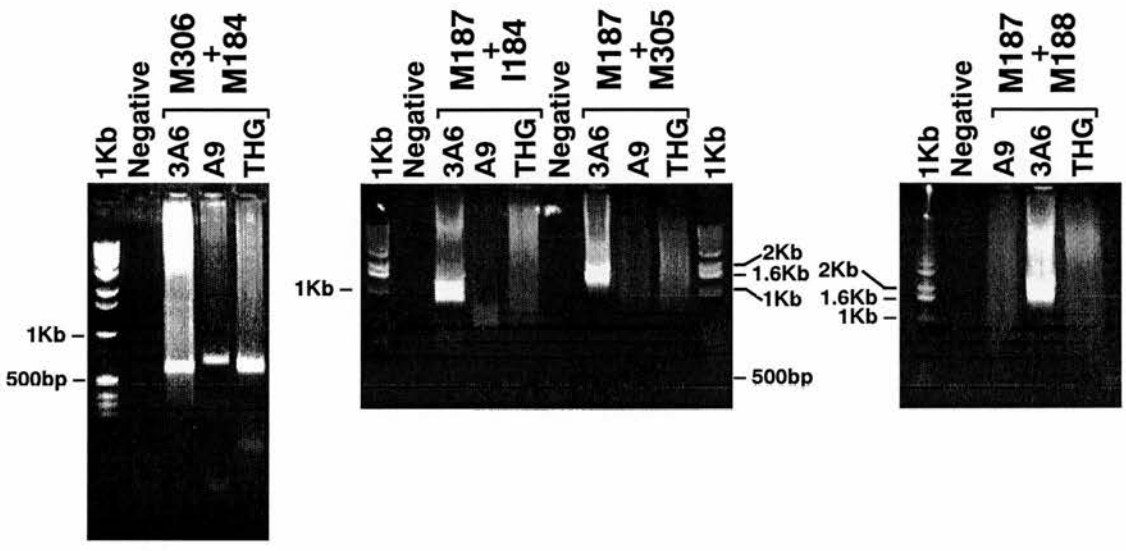


Figure 6.11: PCR analysis showing amplification of fragments across the sequence of the CSC clone 3A6 from genomic DNA (A9 and THG = total human genomic DNA) and 3A6 clone DNA. If the clone 3A6 is not a cDNA then amplification of correctly sized bands across increasingly large fragments of the clone would be expected to occur in both the clone and the genomic DNA. If introns are present in the genomic DNA then amplification should occur in the clone 3A6 but not the genomic DNA.
Expected sizes: M306/M184 - 598bp
M187/I184 - 946bp
M187/M305 - 1501bp
M187/M188 - 1865bp

- Sequencing the CSC clone 3A6;

The clone 3A6 was completely sequenced using the primers described in Table 6.5.

Table 6.5: Primers used to sequence the 2.3Kb CSC clone 3A6

Primer Name	Primer Sequence
M188	5'- CCA GAA GTT TGT CAC CAG -3'
M187	5'- TGA GGC AGG AGA ATC GCT -3'
M305	5'- TTG TCT GCC TGA AAA TGT -3'
M306	5'- CCA GGG ACA AAG TTW WCA -3'
N164	5'- TAA CAA CTA ACC TCC CCT -3'
I184	5'- CTT TCC ACC TTC TGC CAT -3'

W = A or T

Figure 6.10 shows the sequence of CSC clone 3A6 with sequencing primers highlighted in bold. This sequence contains several *Sau3AI* restriction enzyme sites indicating that it is the result of a partial *Sau3A*I digestion. Until this clone is proved to originate from the original cDNA, the possibility that it is a chimaeric clone can not be excluded.

The clone sequence contains three potential predicted poly A sites which could be indicative of the 3' end of a gene. These occur at positions 313bp, 335bp and 1846bp and are represented by the sequence AATAAA.

The demonstration of an open reading frame (ORF) (coding sequence containing no stop or nonsense codons) in a DNA sequence, which can code for a peptide, provides additional evidence of the genic nature of the sequence. The CSC clone 3A6 was therefore translated in all three reading frames in both directions in order to try to demonstrate the presence of an open reading frame. However, the sequence does not maintain a continuous open reading frame throughout its length. The presence or absence of an open reading frame is critically dependent on the sequence quality. A single incorrect base pair can destroy the continuity of the ORF and since there are several ambiguities in the 2.3Kb sequence of this clone this may offer one explanation for the lack of a continuous ORF. Sequencing and cloning errors may have occurred during the CSC procedure which could affected the ORF continuity. The CSC technique is PCR based in which Taq polymerase can incorrectly copy bases. The DNA has also undergone several rounds of cloning (since the DNA used in the CSC experiment had already been cloned into either a

cosmid or PAC) by the time the final CSC DNA clone is sequenced increasing the likelihood of cloning errors which alter the DNA sequence.

The sequence of clone 3A6 contains a large TC repetitive sequence at the 3' end of the sequence. Repetitive sequences identified by repeat database searches ("repeat masker", available from at HGMP resource centre, Hinxton Hall, Cambridge) included a MSTD repeat from the LTR/MaLR family of repeats extending from 1040bp to 1412bp in the sequence, a simple (GGAA)n repeat extending from bases 1875-2013 and an AluSc repeat from the SINE/Alu family extending from 2015bp-2248bp (see Figure 6.10).

The lack of open reading frame and the presence of repetitive elements in the clone sequence are indicative of the sequence comprising the UTR of the corresponding gene. Since the original cDNA resource used in the CSC experiment was oligo dT primed it is likely that 3' sequences of genes will be isolated. However, if the clone sequence does represent a 3' UTR it would be expected that the PCR with primer pairs at either end of the sequence (see previous section and Figure 6.11) should have worked on genomic DNA as well as the cDNA clone DNA, as 3' UTR regions do not usually contain introns.

- Hybridisation of clone 3A6 onto the CSC library:

The 3A6 clone was hybridised back onto the CSC library in order to determine whether any other clones containing this sequence were present. One other clone was positive in this screening. This clone had a very small insert size (41bp). The sequence did not further extend that obtained from the 3A6 clone sequence.

- Mapping 3A6 to somatic cell hybrid panel:

The whole CSC 3A6 clone insert was used as a probe in Southern blot hybridisation analysis of the somatic cell hybrid panel (cut with restriction enzyme *EcoR1*). The results from this hybridisation are displayed in Figure 6.12. Although there is a background smear, present due to the repetitive elements contained in this clone, a distinct band is visible in the chromosome 1 only hybrid A9, in total human genomic DNA and in the derived chromosome 11 hybrid MAR1. This band is approximately 5Kb in size, which is the same size as that seen in the hybridisation of this probe to the *EcoR1* digest of the PAC DJ4B9/3. The band seen in the MAR1 hybrid is weaker than that in total human genomic DNA due to the lower copy number of this

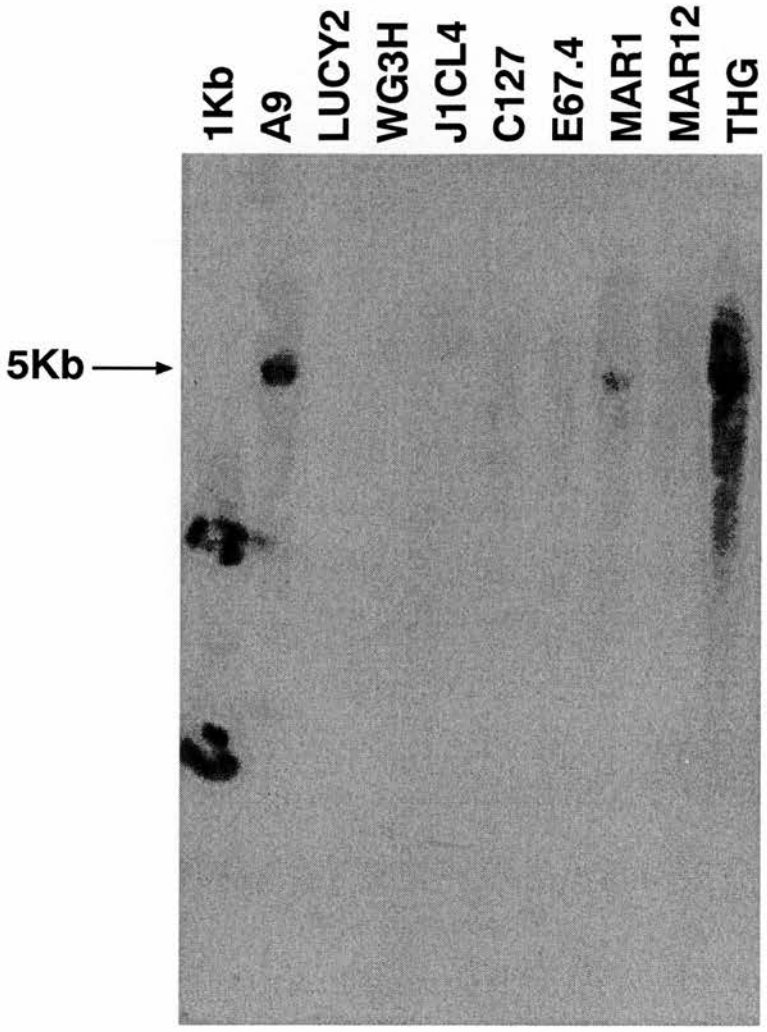


Figure 6.12: Southern blot hybridisation analysis of 3A6 hybridised to somatic cell hybrid panel cut with *EcoRI*

chromosome in this hybrid. This result confirms that the clone maps distal to the chromosome 1 translocation breakpoint.

In conclusion more work on this clone is required to address whether it is truly a coincident product representing a cDNA derived from foetal brain (See discussion). It is still possible that this clone is either a chimaeric clone made up of several unrelated fragments or is a genomic artefact.

Clone 11B6:

- Sequencing clone 11B6:

This clone was identified by random sequencing of 50 CSC clones. The clone insert was 1076bp in length. It was possible to sequence across the whole of this clone using the vector primers 291 and 292. The sequence of clone 11B6 is displayed in Figure 6.13.

Database searching against the GENBANK database showed that this clone matched two ESTs in the database. The first of these ESTs (and the best match - 96% identity over 546bp)) was a foetal lung EST (GENBANK accession number W05481). The second EST clone 11B6 matched was derived from foetal heart DNA (GENBANK accession number AA347756, 92% identity over 436bp). (N.B identity is assessed by the BLAST program (Pearson and Lipman 1988)). The two EST matches extended over the same region of the CSC clone as indicated in Figure 6.14a and b. The foetal lung EST extends over the region 248bp to 498bp and the foetal heart EST extends over the region 231bp to 433bp. The foetal lung EST was derived from an I.M.A.G.E consortium clone which was obtained from the HGMP resource centre, Hinxton Hall, Cambridge. The size of this clone was approximately 600-700bp as determined by a PCR reaction carried out using primers designed to the vector (pT3T7D) sequence (T3 5'- AATTAACCCTCACTAAAGGG -3' and T7 5'- CATTATGCTGAGTGATATCCCG -3'). This foetal lung clone, named 299277, was sequenced using the vector primers. However despite several attempts the clone was refractory to sequencing.

Figure 6.13: Sequence of CSC clone 11B6. Capture oligonucleotide sequence is highlighted in bold and repetitive elements are marked in the sequence.

```
   1  GAATTCAGAC ACTGGTATGC ATCTCTATTA ATGTGTAATA AGTGGCTCCT
  51  CAAACTTCAT AGCTTGAAAC AACAAACCTT TATCTCTCAC AGTCTCTGAG
 101  AGACAGGAGT CCAGAGTGAC TTAGCTGGGT GGCTCTGACT CAGGGTCTCA
 151  TAAGGTCGCT GTTAAGACAC ACTTTGGAGC TGCAGTCGTC TAAAGGCTTG
 201  GCTGGGGCTG AAAGATCACA GAGGCATGGG CACTGTGTTC TAAGTGGTCA
 251  CTGCACTGAT TCAGAACAGC AGGGGCTGGC TCTGTACTGG GTGGGTGGGT
 301  GCTCAAGGCC AGACCTACAC AGTGCTCCTG TGTCTGCAGC TGCAGGGAAG
 351  CAGGAGAAAC AGTGACGATG GTCCRGGAGA AGCCACCTGA CATATGGCAG
 401  AAAAACAAAA TTCAGGGTAC AGACAGTGGS TGGGAVCATC AAYTCAATYG
 451  CTTTCTTTCT TTAMTATTTT CCYTCTTTCT AAAAAAGTCT GTGATTTAAA
 501  TCAGTGGCTG CAGGARGGAA ACAAGARAMC CAGTAAAGAT GTTTTCAAAG
 551  ATGATCCTAT CTGGTGTGAA AAGAAATGAA RAMCTACCCA AAGGGAATAA
 601  GCACAAGTAT TATGCAGAGC TTKGTACAGC AAGGGAKTCA GACACCATCA
 651  TTTGCAATTT GGCAGAGATT CAAAGGTGGG CAGACGAGTG GGAGAGCTTT
 701  ATAGTGGAAA AAGGCGAAGG CTTCAGGTGT GCCCTGATTG GAGGTTATCA
 751  ATGCGGGGAA GCTGGAGGCG GCTCACTAGA AGCGAACATC CTGTGTGCTT
 801  GGTCAGGGGA CCATATTTGG CTTTCTCTGG TGGGTCCTAA GTTGGAATCG
 851  GGGACAAAAA TTAGGGAAGC CATCAGTTAT TAATCCAGTC CTGAACATTT
 901  TGAGTCAATT GTTACAGAAG TTATTATTTA GCTTCCTGGA TAGTTACTAG
 951  AGAGCAATTT GGCTTCCAGC AGGTCTGATT TAGAGCAGGC CAGCTTCCGG
1001  GGTTGCTTTT TGTGGGTAAG GGTATTGTTT TCTGGGAAAG TTGCTGCACC
1051  TTGTGGATCG GTCGACTCTA GAATTC
```

MLT1F REPEAT

K= T/G
M= A/C
R= A/G
S= C/G
W= A/T
Y= T/C

Figure 6.14a: Sequence alignments of CSC clone 11B6 and two ESTs identified as being of strong identity to this clone by GENBANK database searching.

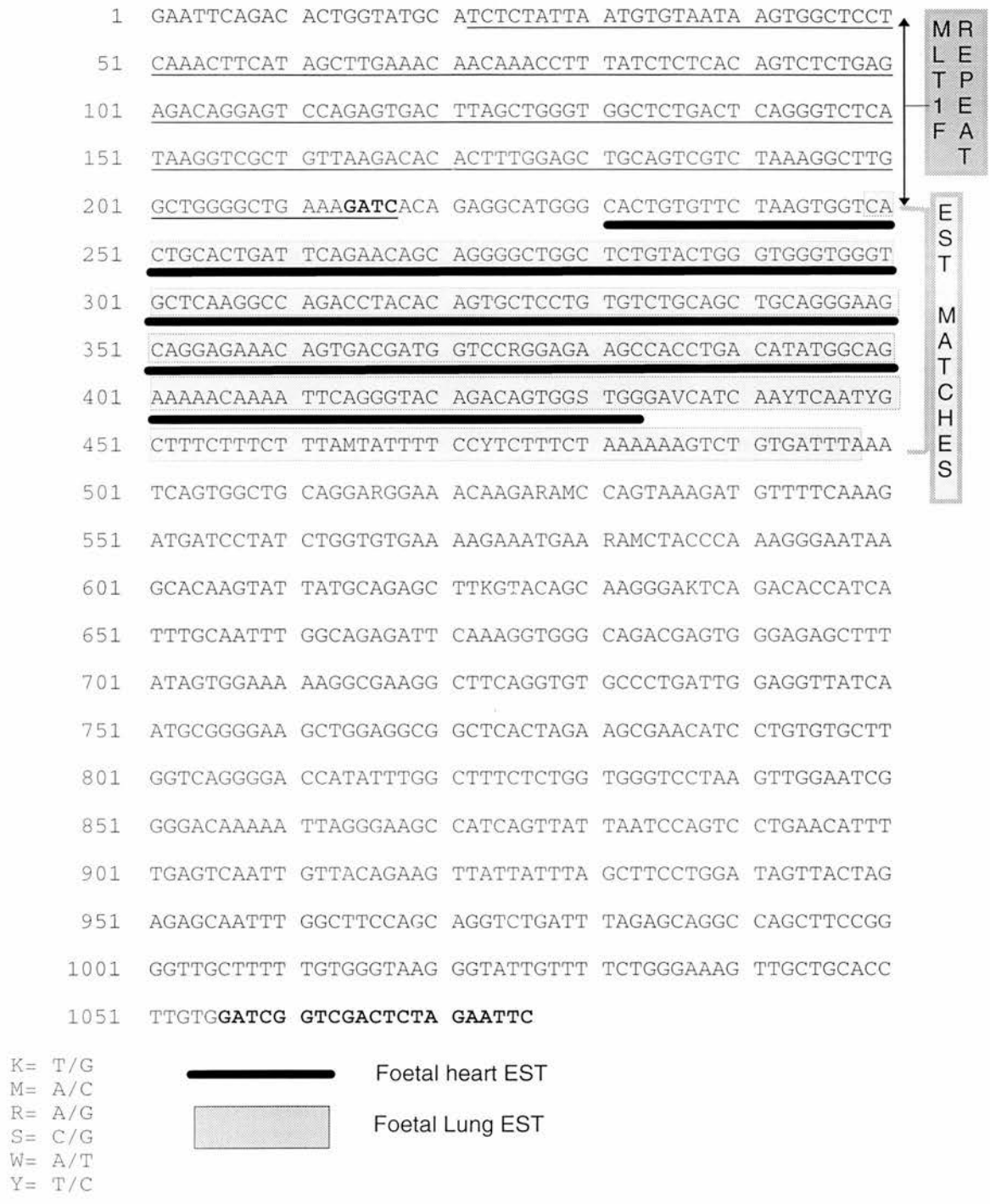
I: Sequence alignment of CSC clone 11B6 with a foetal lung EST identified from GENBANK database searching.

```
GB:W05481 W05481 ZA84F11.R1 SOARES FOETAL LUNG NBHL19W HOMO
          SAPIENS CDNA CLONE 299277 5'.
          LENGTH = 546


IDENTITY= 96%


11B6
CACTGCACTGATTCAGAACAGCAGGGGCTGGCTCTGTACTGGGTGGGTGGGTGCTCAAGG 308
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTGCACTGATTCAGAACAGCAGGGGCTGGCTCTGTACTGGGTGGGTGGGTGCTCAAGG 275
EST

11B6
CCAGACCTACACAGTGCTCCTGTGTCTGCAGCTGCAGGGAAGCAGGAGAAACAGTGACGA 368
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAGACCTACACAGTGCTCCTGTGTCTGCAGCTGCAGGGAAGCAGGAGAAACAGTGACGA 215
EST

11B6
TGGTCCRGGAGAAGCCACCTGACATATGGCAGAAAAACAAAATTCAGGGTACAGACAGTG 428
|||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
TGGTCCAGGAGAGGCCACCTGACATATGGCAGAAAAACAAAATTCAGGGTACAGACAGTG 155
EST

11B6
GSTGGGAVCATCAAYTCAATYGCTTTCTTTCTTTAMTATTTTCCYTCTTTCTAAAAAAGT 488
| ||||| |||||| ||||| |||||||||||||| |||||||| |||||||||||||||
GCTGGGAGCATCAACTCAATCGCTTTCTTTCTTTACTATTTTCCCTCTTTCTAAAAAAGT 95
EST

11B6 489 CTGTGATTTA 498
         ||||||||||
EST  94  CTGTGATTTA 85
```

II: Sequence alignment of CSC clone 11B6 with foetal heart EST identified from GENBANK database searching.

```
GB:AA347756 AA347756 EST54107 FOETAL HEART II HOMO SAPIENS CDNA 3'
          END.
          LENGTH = 436
IDENTITY = 92%

11B6
CACTGCACTGATTCAGAACAGCAGGGGCTGGCTCTGTACTGGGTGGGTGGGTGCTCAAGG 308
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTGCACTGATTCAGAACAGCAGGGGCTGGCTCTGTACTGGGTGGGTGGGTGCTCAAGG 285
EST

11B6
CCAGACCTACACAGTGCTCCTGTGTCTGCAGCTGCAGGGAAGCAGGAGAAA-CAGTGAC- 366
|||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
CCAGACCTACACAGTGCTCCTGTGTCTGCAGCTGCAGGGAAGCAGGAGAAACCAGTGACG 345
EST

11B6
GAT-GGTCCRGGAGA-AGCCACC-TGACATA-TGGCAGAAAAA-CAAAATTCAGGGTACA 421
||| ||||| || ||  |||||| ||||||| |||||||||| ||||||||||||||||
GATNGGTCCNGGNGAGGGCCACCTTGACATATTGGCAGAAAAACCAAAATTCAGGGTACA 405
EST

11B6 422 GACAGTGGS-TGGGAVCATCAAYTCAAT 448
         | ||||||  ||||  |||||| |||||
EST  406 GNCAGTGGCTTGGGNGCATCAACTCAAT 433
```

Figure 6.14b: Sequence of CSC clone 11B6 indicating region to which the ESTs identified by database searching match. Capture oligonucleotide sequence is highlighted in bold and repetitive elements are marked in the sequence. EST matches from GENBANK database are also displayed.



Repeat database searches with the sequence of the clone 11B6 showed that this clone contained a repetitive element. This repetitive element was a MLT1F repeat which is a member of the LTR/MaLR family of repeats. The repeat extends from position 22bp to 217bp, where there is *Sau3AI* restriction enzyme site. These repetitive elements are very low copy in the human genome. This clone was not identified as a repeat containing clone in the initial library screen with Cot1 DNA. This is likely to be due to under representation of this repeat in the Cot1 DNA which is largely composed of Alu repeats.

The sequence of 11B6 was checked for the presence of open reading frames which would help indicate the genic nature of this DNA fragment by translating the sequence in all three frames in both directions. The sequence does not contain any sustained open reading frame. This may be for several reasons which have been mentioned previously for clone 3A6. It may be that errors from the sequencing of the clone prevent the maintenance of an open reading frame or that the sequence represents 3' untranslated region of the corresponding gene and no coding sequence is present in this clone. There are no other distinguishing features in this sequence (such as a consensus poly A site) which could define its genic nature further. The fact that the sequence contains a repetitive element may suggest that it is UTR sequence. It is possible that this clone is chimaeric. The match of this clone to the two ESTs in the database begins abruptly with a very good match at a *Sau3AI* restriction enzyme site (GATC)(refer to figure 6.14b; *Sau3AI* sites highlighted in bold). Prior to this *Sau3AI* site, despite the fact the EST sequence continues, there is no reasonable match to the sequence possibly indicating that this clone is chimaeric. The repeat containing section up to the *Sau3AI* site being unrelated to the region of the two EST matches and having arisen from an aberrant ligation reaction.

- Is clone 11B6 derived from the genomic resource?

As previously mentioned for clone 3A6, it is important to determine that the clone has been derived from both the genomic and cDNA resources that were originally used in the CSC experiment. The 11B6 clone insert was therefore used as a probe in Southern blot hybridisation to the same PAC and cosmid digest filters as with clone 3A6. The results from this hybridisation are displayed in Figure 6.15. It is apparent from this result that bands can be seen in the cosmid I0142 and in the PAC DJ4B9/3, but not in the cosmid B01519. In the *BamH1* digest of the PAC and

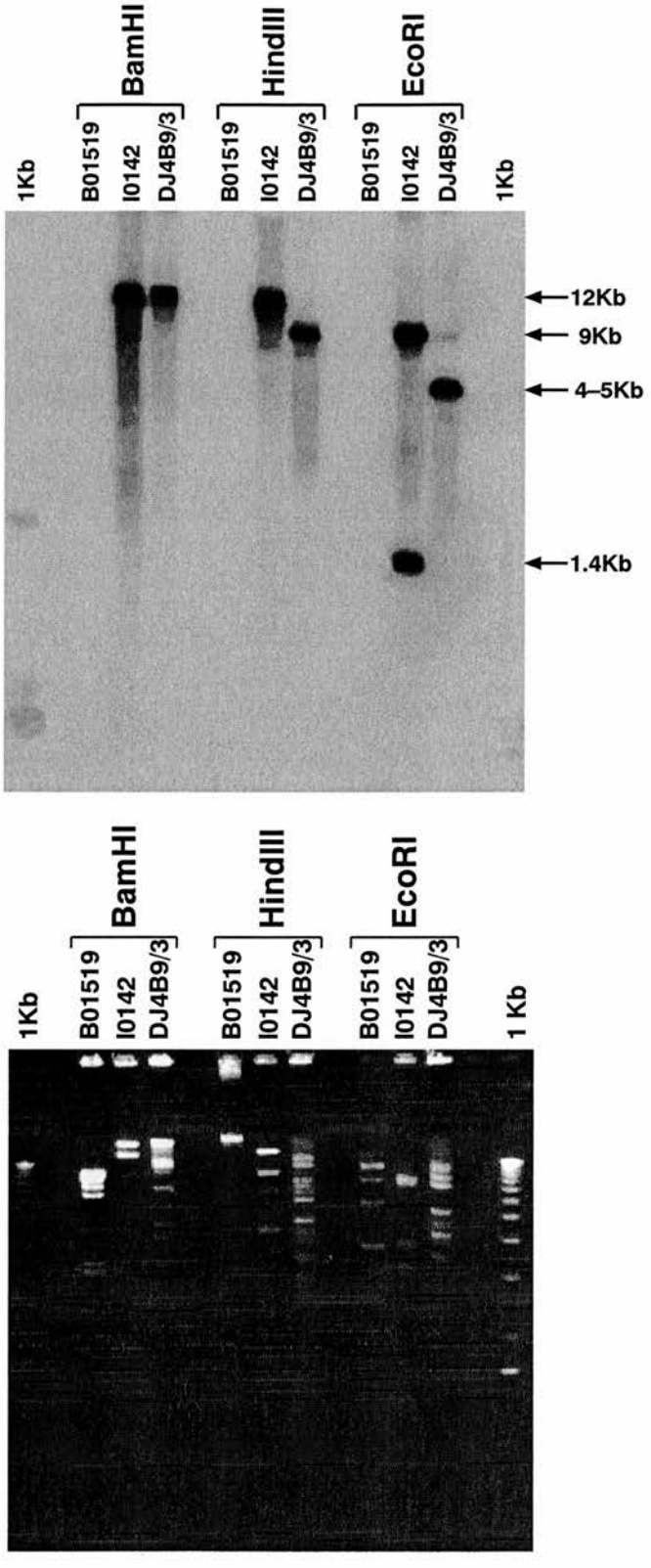


Figure 6.15: Southern blot hybridisation of CSC clone 11B6 hybridised to restriction enzyme digests of the cosmids BO1519 and I0142 and PAC DJ4B9/3.

cosmids two bands are visible in very close proximity to one another at approximately 12Kb. The *HindIII* and *EcoRI* digests of the cosmids and PAC also have bands in the PAC and I0142 cosmid lanes but these bands are of different size (see Figure 6.15). The bands seen in the PAC are of smaller size than those in the cosmid I0142 in these two digests (*EcoRI* and *HindIII*). This is due to the cosmid I0142 extending beyond the PAC distally in the contig and therefore a smaller restriction fragment may be expected to be seen in PAC as the clone maps at the extreme end of the cosmid reducing the size of the restriction fragment. There is a background smear in the positive lanes (PAC and cosmid I0142) which is attributable to the presence of the repetitive element which was identified from sequence analysis of this clone.

- Is the CSC clone 11B6 derived from foetal brain cDNA ?

In order to determine if this clone was derived from foetal brain a PCR on the original cDNA was carried out using primers designed to the clone sequence (N811: 5'-GCA GGG AAG CAG GAG AAA -3'; N812: 5'- TCC CAG AAA ACA ATA CCC -3'). The PCR was achieved using A9 DNA (human chromosome 1 only), 11B6 minipreped DNA and original cDNA as template. PCR conditions are described in Table 6.6. PCR products of the correct size were produced from the chromosome 1 only DNA and from the 11B6 clone DNA but not from the cDNA originally used in the experiment (see figure 6.16). This may be due to the CSC clone not being highly represented in this DNA.

Table 6.6 PCR condition for amplification of CSC clone from original cDNA resource.

	Denaturing	Annealing	Extension
Cycle 1	94°C/30sec	50°C/45sec	72°C/60sec
Cycle 2-34	94°C/15sec	50°C/45sec	72°C/60sec
Cycle 35	94°C/15sec	50°C/45sec	72°C/90sec

In order to demonstrate that the clone was derived from foetal brain cDNA the whole clone insert was used to screen a foetal brain cDNA library (Clontech). A

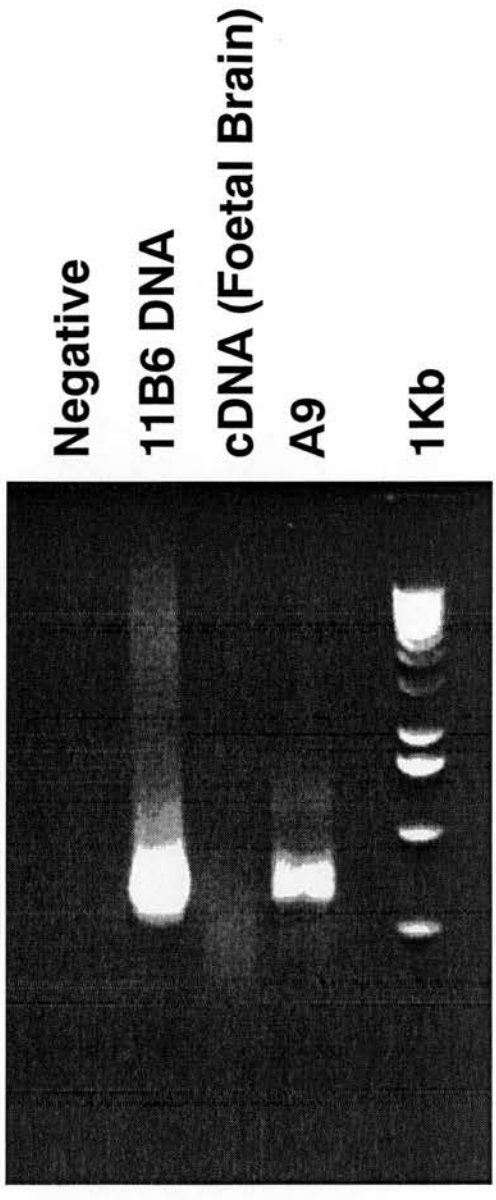
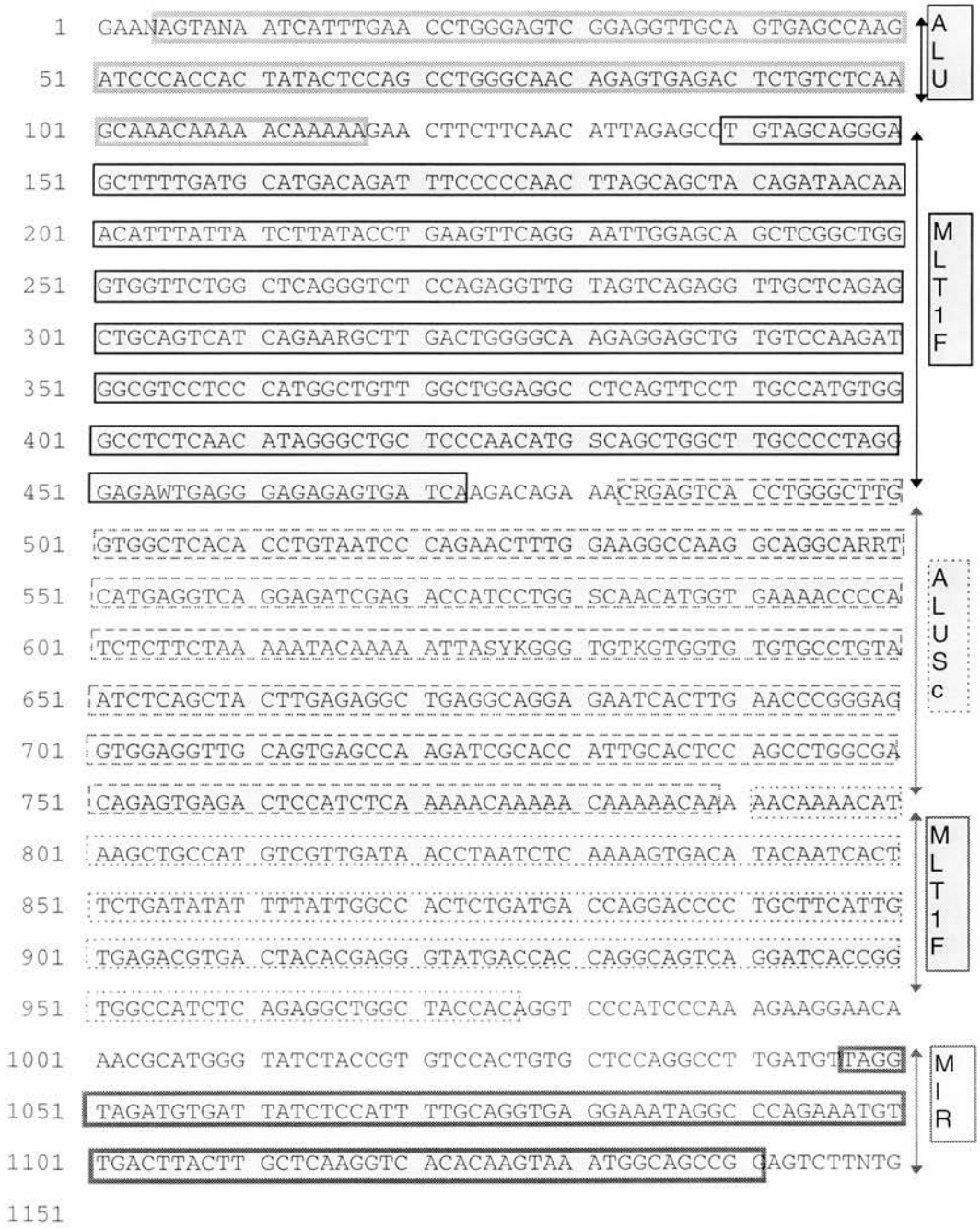


Figure 6.16: PCR analysis showing amplification of 11B6 from original foetal brain cDNA resource, CSC clone DNA and Human chromosome 1 DNA (A9).

number of positive clones were obtained from this screening. The inserts were amplified by boiled plaque PCR and sequenced using the vector primers 391 and 392. The sequence obtained from these cDNA clones is presented in Figure 6.17A. When the sequence of cDNA clone 311 was compared to that of the CSC clone

Figure 6.17: cDNA clone sequences obtained by screening a full length foetal brain cDNA library with CSC clone 11B6.

A) Sequence of cDNA clone 31. Repetitive elements are highlighted.



B) Sequence comparison of CSC clone 11B6 and cDNA clone 31 obtained from screening a foetal brain cDNA library with the CSC clone 11B6. There is only 64% identity despite stringent hybridisation conditions.

```
            64.2% identity in 193 bp overlap

                    10        20        30        40        50
11B6          GAATTCAGACACTGGTATGCATCTCTATTAATGTGTAATAAGTGGCTCCTCAAA
                                    || || |||  | | |  |   ||| | ||
cDNA    TTCTTCAACATTAGAGCCTGTAGCAGGGAGCTTTTGATGCATGACAGAT--TTCCCCCAA
              130       140       150       160       170

            60        70        80        90       100       110
11B6    CTTCATAGCTTGAAACAACAAACCTTTATCTCTCACAGTCTCTGAGAGACAGGAGTCCAG
        |||   ||||  | | |||||||| ||||| | ||  | |  ||||    ||||| |   |
cDNA    CTTAGCAGCTACAGATAACAAACATTTAT-TATCTTA-TACCTGAAGTTCAGGAAT-TGG
      180       190       200       210       220       230

           120       130       140       150       160       170
11B6    AGTGACTTAGCTGGGTGGCTCTGACTCAGGGTCTCATAAGGTCGCTGTTAAGACACACTT
        ||   ||  ||||||||| |||| ||||||||||   |||| | |  | |||      |
cDNA    AGCAGCTCGGCTGGGTGGTTCTGGCTCAGGGTCTCCAGAGGTTG-TAGTCAGAGGTTGCT
         240       250       260       270       280       290

           180       190       200       210       220       230
11B6    TGGAGCTGCAGTCGTCTAAAGGCTTGGCTGGGGCTGAAAGATCACAGAGGCATGGGCACT
          |||||||||| ||  ||:|||| |||||||   | || |
cDNA   CAGAGCTGCAGTCATCAGAARGCTTGACTGGGGCAAGAGGAGCTGTGTCCAAGATGGCGT
          300       310       320       330       340       350

           240       250       260       270       280       290
11B6    GTGTTCTAAGTGGTCACTGCACTGATTCAGAACAGCAGGGGCTGGCTCTGTACTGGGTGG

cDNA    CCTCCCATGGCTGTTGGCTGGAGGCCTCAGTTCCTTGCCATGTGGGCCTCTCAACATAGG
          360       370       380       390       400       410
```

11B6 the identity was only 64% (Figure 6.17B). Since the hybridisation and washing conditions had been stringent (68^{0}C, 0.1X SSC) this was surprising. Database searches against the GENBANK and "repeat masker" databases (HGMP, Hinxton Hall, Cambridge) showed that in fact the sequence of the cDNA clone 311 was largely composed of repetitive elements (see Figure 6.17A).

The CSC clone 11B6 used to screen the cDNA library contains a repetitive element. However, as the cDNA library screening was carried out prior to the full sequence of the clone being known, this repetitive element was not blocked prior to the library screening. The clone from the cDNA library has been identified due to homology with this repetitive element. This CSC clone has not therefore been proved to be a cDNA by this library screening.

A foetal heart cDNA library has also been screened with this clone (in which the repeat has been competed out) and has resulted in several strong positives which are currently being reduced to single plaques.

- Mapping the clone 11B6 to the somatic cell hybrid panel:

It is not possible to tell from the hybridisation to digests of the PAC and cosmids whether the CSC clone 11B6 maps proximal or distal to the chromosome 1 translocation breakpoint. In order to determine this, the clone insert was hybridised in Southern blot hybridisation analysis to a panel of somatic cell hybrids containing the chromosome 1 only hybrid A9, total human genomic DNA and the translocation hybrids MAR1 and MAR12. The result of this hybridisation is displayed in Figure 6.18. A band at approximately 7Kb is present in the chromosome 1 only hybrid A9, in total human genomic DNA and in MAR1. This indicates that the clone 11B6 lies distal to the chromosome 1 breakpoint. In the total human genomic lane several bands can be seen. This is probably due to the repetitive element in the clone but also may indicate that this clone has other related genes or pseudogenes present elsewhere in the genome since not all of these extra bands are not seen in the chromosome 1 only hybrid A9.

- Hybridisation of CSC clone 11B6 to the CSC library:

The whole 11B6 clone insert was also hybridised onto the CSC library in order to determine if any other related clones were contained in the library. No other positives clones were identified during this screening.

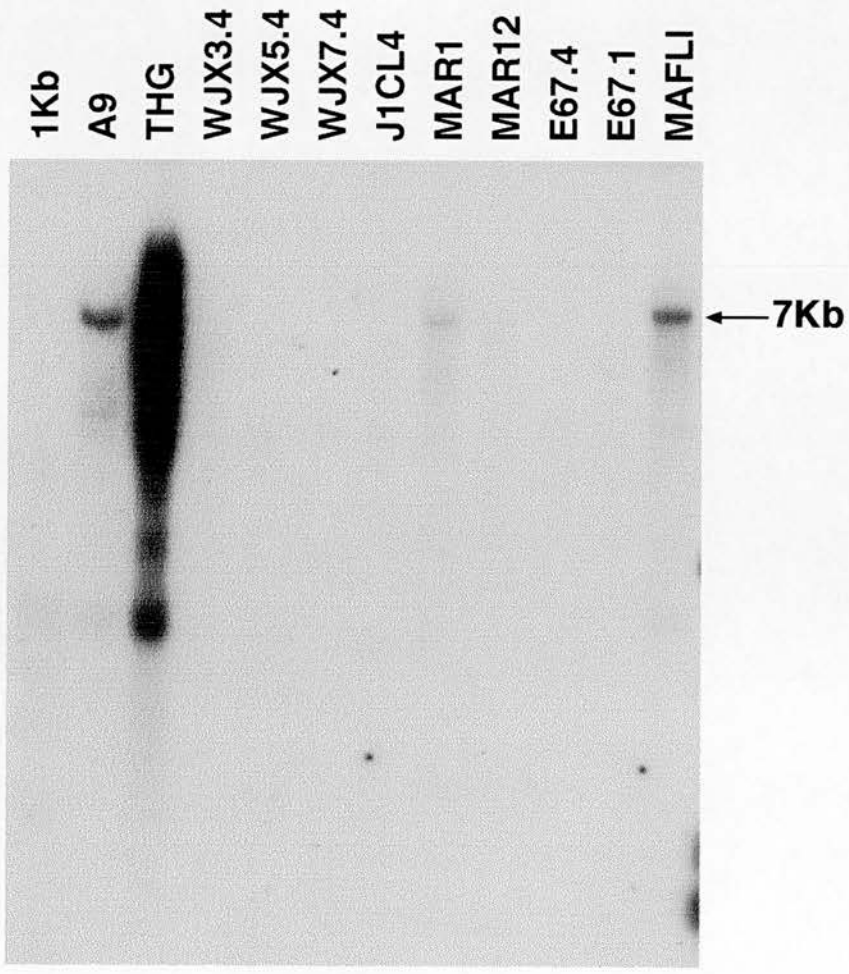


Figure 6.18: Southern blot hybridisation of clone 11B6 hybridised to the somatic cell hybrid panel cut with *EcoRI*. (THG= Total human genomic DNA).

6.5 Discussion:

Several gene fragments have been identified as residing around the region of the translocation breakpoint using several differing methods. These gene fragments are currently being investigated (ongoing work by Kirsty Millar, Susan Anderson and Sheila Christie. See section 7.2)). The Human Transcript Map (Schuler et al 1996) has allowed further gene fragments residing around this region to be isolated. This region of chromosome 1 appears to be gene rich since the Human Transcript Map identified approximately 60 genes in a 1Mb region around the translocation breakpoint.

The average density of genes in the human genome is 1 every 40Kb. The region around the chromosome 1 breakpoint appears to be very gene rich in comparison to the chromosome 11 breakpoint around which there appears to be a paucity of genes. Many of the gene fragments isolated around the chromosome 1 breakpoint may turn out to be parts of the same gene or constitute several genes residing in the region. The region considered in this experiment was approximately 140Kb in size and would therefore be expected to contain 3-4 genes on this basis. If the two DNA clones identified prove to be real gene fragments then despite the severe contamination problems encountered the experiment can be counted as successful. The two DNA clones will be further investigated to determine their candidacy as causal elements involved in the psychiatric diagnosis in the K26 translocation pedigree.

Several criteria were used in this experiment to assess the products which were obtained from the CSC experiment. These criteria are listed in Table 6.7

Table 6.7: Criteria for assessing cloned products derived from CSC experiment

CRITERIA
Clone must be derived from genomic resource starting material
Clone must be derived from original cDNA starting resource
Clone must be shown to be of human origin by mapping to total human genomic DNA and A9 chromosome 1 only hybrid
Clone sequence should be compared to sequences in the GENBANK, dbEST and repeat databases to determine if it is novel or of known sequence and it does not consist entirely of repetitive

elements or vector sequences
Clones should be correctly linkered and contain the correct vector sequence at the beginning and end of the sequence

The two clones identified in this experiment have met all but one of these criteria. Neither of the two clones have been shown to be derived from the original cDNA.
In the case of clone 11B6 the CSC clone sequence matches that of two ESTs in the database. Whereas this does not prove that it is a cDNA it does show this clone to be interesting, even if it does correspond to a chimaeric/genomic clone, since it maps to a region close to the breakpoint (if it is a chimaeric clone then one segment contains the EST match which is of interest and the other smaller segment contains repetitive sequence). The EST matches are derived from two different libraries. This would imply that the EST matches are not coincidental matches resulting from a mutual contaminant occurring in the CSC library and the cDNA library from which the ESTs were derived. This clone has also given strong positive results when used as a probe onto a foetal heart library (the same tissue from which one of the EST matches was derived). Although sequence has yet to be obtained from these positive clones from this library screening this provides further evidence that this clone contains expressed sequence which resides in close proximity to the breakpoint and it is therefore of further interest.

The 3A6 clone does not have any evidence of containing expressed sequences and must be shown to be derived from foetal brain cDNA (or be genic in nature) in order to be of further interest. It is likely that this clone is either chimaeric or a genomic artefact due to its large size and multiple *Sau3AI* restriction sites. The end ligation procedure is very stringent and in order to be a genuine product this clone would have to have hybridised to a cDNA of similar size which would also have to have been partially digested in order for the capture oligonucleotides to be ligated properly. If the clone is either chimaeric or a genomic artefact, it may be difficult to assess the true nature of this clone.
The CSC technique is sensitive to contamination from various sources due to its amplification stages which amplify the contaminant as well as genuine products. All such contamination is problematic, as has been clearly demonstrated in these experiments but particularly so if it is derived from the original genomic resource,

emphasising the importance of demonstrating that the clones are genic in nature and derived from both the original starting resources.

Although the CSC experiment recovered two novel products from the region around the translocation breakpoint on chromosome 1 the high contamination level of the library was undoubtedly detrimental to the isolation of further transcripts. It would be appropriate to repeat the CSC experiment either using PAC DNA of a purer quality, which could be achieved by further gel purification steps to remove contaminating *E.coli* sequences, or just using the cosmid DNA as the genomic resource. This may circumvent the severe contamination problems experienced during the course of these experiments. The high level of contamination of the library was also probably the reason that the hybridisation with DNA probes known to map to the region (i.e. the 7Kb wild type chromosome 1 fragment spanning the breakpoint and the EST probe which lies within the PAC) did not result in the isolation of any positive clones.

It would also be worth persevering to resolve the non-specific binding of the cDNA resource to the streptavidin coated magnetic beads experienced in the hybrid fishing technique, since the end ligation method, although more stringent, favours the isolation of intra exonic fragments which will fail to recover some genuinely coincident products. The combination of the hybrid fishing and the end ligation methods would result in more gene fragments being isolated.

The most significant drawback of the coincident sequence cloning technique is its limitation to only identify genes expressed in the tissue chosen to make the cDNA resource. The selection of the tissue in which the causative gene will be expressed is therefore critical to the procedure. Similarly, the developmental stage of the tissue chosen may be critical for the gene to be expressed. In this experiment using foetal brain tissue from three stages of development was aimed at circumventing these problems as far as possible, since foetal brain expresses a large percentage of the genes of the genome (>50%) and a gene involved in psychiatric illness would most likely be expressed in the brain. However, complementary gene isolation techniques should also be considered such as exon trapping or direct screening of cDNA libraries. cDNA selection appears to be the preferred method for gene isolation in the literature and Harshmann et al (1995) showed that this method was more efficient than direct library screening or exon trapping contributing to 39/45 candidate gene fragments isolated from the BRCA1 region on chromosome 17q.

The CSC experiment has allowed the isolation of two novel clones in close proximity to the chromosome 1 translocation breakpoint. These clones can now be further investigated as potential candidate genes involved in the psychiatric diagnosis in the K26 translocation pedigree.

Further investigation of these clones to assess their candidacy would firstly need to confirm that the clones were derived from foetal brain cDNA (especially in the case of 3A6). This could be achieved by screening foetal brain cDNA libraries with different probes derived from the clone sequence. In the case of 3A6 this may be important if this clone is chimaeric as only a small portion of the clone may contain the cDNA fragment.

Once it has been established that the clones are derived from foetal brain cDNA, the sequences should be extended to full length by repeated screening of full length foetal brain cDNA libraries(or libraries derived from other tissues). Sequences which are not extended fully by this procedure could be extended using the technique for 5' or 3' **R**apid **A**mplification of **c**DNA **E**nds (RACE).

The information gained from the additional sequence obtained using these techniques will allow further database searching to identify matches to known genes at the nucleotide and protein level. Analysis of the predicted proteins structure by hydrophobicity profiling and identification of protein domains (such as membrane spanning domains) may provide clues as to the function of the gene products corresponding to these clones.

The clone sequences once extended should be mapped at high resolution with respect to the translocation breakpoint. This could be achieved by restriction enzyme mapping of the PAC and cosmids from which the clones were derived. Determining the proximity of these clones to the breakpoint in combination with extension of the sequences to full length will indicate how far these clones extend towards the breakpoint and if either of them cross the breakpoint. Several other gene fragments have been identified in close proximity to the breakpoint (work carried out by Kirsty Millar, Susan Anderson and Sheila Christie). These sequences may be related to, or part of the same gene as the CSC clones. Fine mapping in relation to these other gene fragments is therefore desirable.

Information on the genomic structure of the gene and its control elements would also be useful to determine if any of these elements are disrupted by the translocation event should the genes prove to be in very close proximity to the breakpoint. The genomic structure of the gene would also be useful to know in

order to characterise intron/exon boundaries. Comparison of the gene in normals and a translocation carrier could also be important in determining differences in the gene structure due to the translocation event.

Analysis of the encoded proteins themselves may reveal information about the function of the genes. Western blot analysis could determine the size of the protein in different tissues and techniques such as immunofluorescence may reveal the localisation of the protein within the cell. Co-immunoprecipitation experiments and the yeast two hybrid system would reveal protein-protein interactions.

Expression analysis by Northern blot, RT-PCR and RNA *in situ* hybridisation techniques would be valuable to determine the expression pattern of the genes relating to these clones. Analysis in different tissues and developmental ages could suggest a possible functional role for these genes. Northern blot analysis would also provide an indication of the size of the genes corresponding to these clones.

In the longer term, if these clones prove to be interesting candidate genes, the elucidation of homologous genes in other species such as the mouse (assessed by zoo blots initially and then isolation of homologous genes by screening a mouse library) would be valuable. These homologous genes may be of known function helping to further elucidate the role of the genes in the pathogenesis of the psychiatric disorders seen in the translocation family.

Transgenic animal models will undoubtedly play a central role in the determination of the effect of the gene in vivo. Gene knock-outs may establish the importance of the genes role in development and manipulations of the gene such as the effects of point mutations, partial deletions or over-expression can be tested to determine their effect by assessing the resultant phenotype. Insertion of a reporter gene such as β galactosidase can be employed to follow the expression of the gene through developmental stages.

Chapter Seven

Discussion

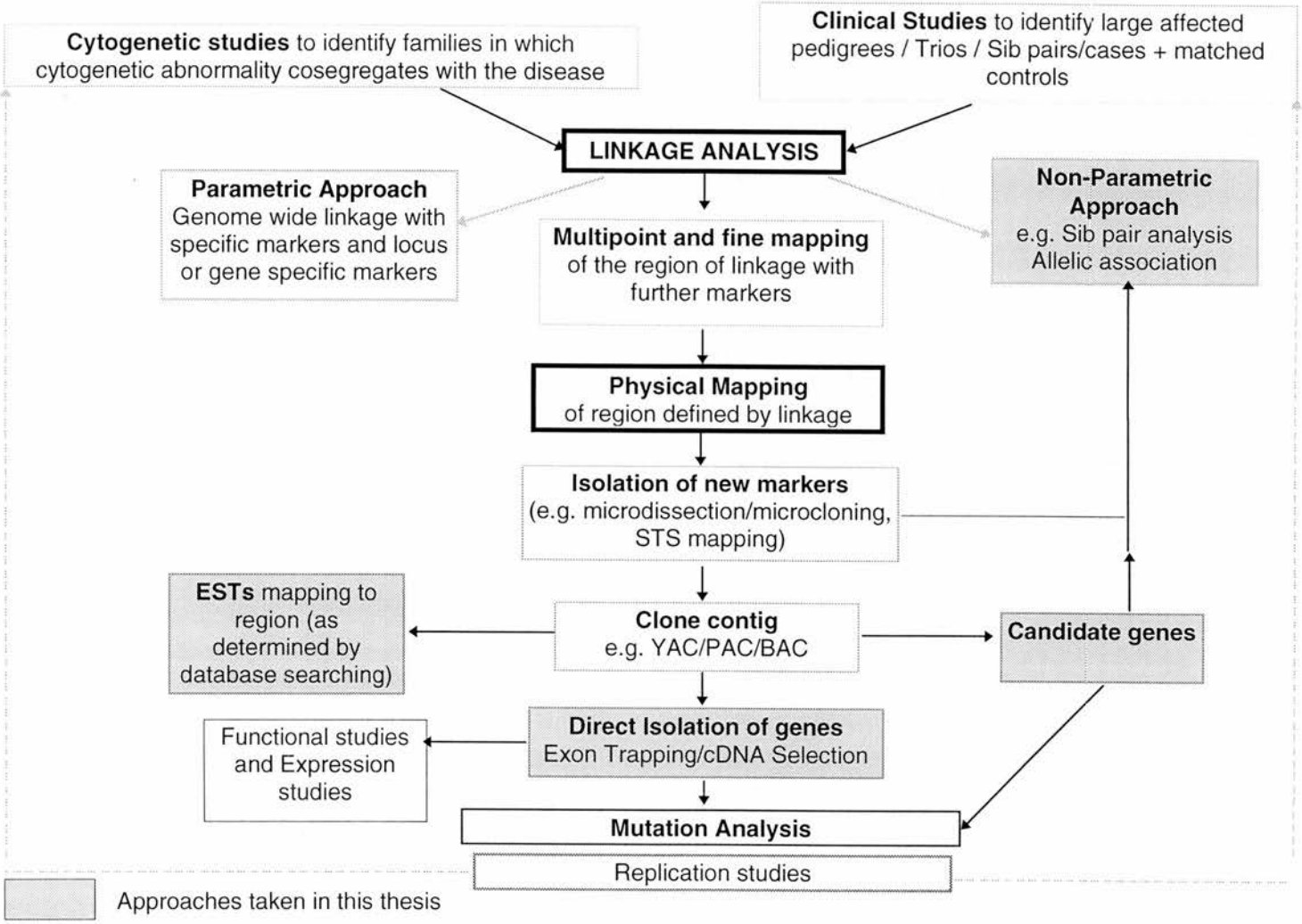
Discussion

Having identified a family carrying a balanced translocation between chromosome 1 and 11 which cosegregates with mental illness, a positional cloning strategy has been employed in order to elucidate the relationship between the translocation and the psychiatric diagnosis in this family. This positional cloning strategy is based on the hypothesis that a gene or genes involved in mental illness reside at or near the translocation breakpoint on chromosome 1 and/or chromosome 11.

A schema of a general positional cloning strategy is illustrated in Figure 7.1. This schema indicates the many sequential stages which can be taken when employing such a positional cloning strategy. Ultimately, despite the route taken, the final outcome should be the identification of disease associated gene(s). The work described in this thesis has used four of the approaches outlined in this schema (these are highlighted in the figure). Although different, these complementary approaches have been employed to identify genes which may be involved in the psychiatric diagnosis in the translocation family.

One of the most obvious and potentially productive ways to determine genes involved with disease processes, is to look initially at those genes already known to reside in the region to which linkage has been found, or which reside close to the cytogenetic abnormality of interest, as in our case. These genes may be determined as candidates due to their known biological function or, initially, solely due to their close proximity to a chromosomal abnormality. The first approach taken in this thesis, to determine genes involved in the psychiatric diagnosis in the t(1:11) translocation family, was such a candidate gene approach. Two α-actinin genes, ACTN 2 and ACTN 3, were considered. These two genes were known to reside in reasonable proximity to the translocation breakpoint and were of additional interest as they were located on both chromosome 1 and 11. High resolution mapping of these genes employing somatic cell hybrids and YACs from the region, allowed these genes to be excluded from further consideration based on their position out with the YAC contigs and their relatively remote location from the translocation breakpoint. Alternative approaches were pursued to determine other candidate genes residing in the region of the translocation breakpoint.

Figure 7.1 Schema of general positional cloning strategy to isolate disease genes.



The Human Genome Project has greatly aided positional cloning. One of the aims of the Human Genome Project is to produce a human transcript map by the year 2005. Towards this aim, a great number of expressed sequence tags (ESTs), representing the ends of cDNA clones derived from various tissues, have been generated and are available in a number of databases including GENBANK and dbEST. Since ESTs are derived from the ends of cDNA clones they should represent gene transcripts and provide an excellent resource for determining the mapping positions of these gene transcripts. The sequence of many thousands of ESTs are now available in several databases. The dbEST database currently contains 819,082 ESTs (September 1997). It is however unclear what percentage of the total number of the estimated 100,000 genes in the human genome this number truly represents as the level of redundancy within this database is uncertain Although many thousands of ESTs are available in the database remarkably few of these ESTs have been assigned to precise mapping locations to date, limiting their immediate usefulness. This situation will undoubtedly change as the Human Transcript Mapping phase of the Human Genome Project picks up speed and moves towards completion of the Human Transcript Map.

The second approach taken in this thesis was high resolution mapping of a number of ESTs which had been described as mapping to the general region of the chromosome 11 translocation breakpoint by Rosier et al (1995). These experiments utilised somatic cell hybrids and YACs from the region and allowed the location of the ESTs in relation to the translocation breakpoint on chromosome 1 to be determined. All of the ESTs considered were excluded from further study due to their remote location from the translocation breakpoint on chromosome 11. However, this mapping approach highlighted difficulties that the presence of repetitive elements can impose (see chapter four). The generation of ESTs from the 3'UTRs of genes is beneficial in that these regions tend to be less conserved than coding regions of genes making it easier to distinguish between individual genes and paralagous gene family members that may be closely related in coding sequence which is desirable for mapping purposes. However, the 3'UTR sequences often contain repeat elements which can cause difficulties when trying to map the precise location of these ESTs as highlighted in chapter four. The use of 3'UTR generated ESTs is less desirable for database searching and sequence analysis as

there is often little coding sequence data and makes extension of these sequences into full length cDNA representing a gene more challenging.

The third approach described in this thesis was an allelic association study with polymorphic markers residing in close proximity to the translocation breakpoint on both chromosome 1 and 11. The allele frequencies of these markers in affected and control populations were compared. Two affected populations were considered; schizophrenia subjects and unipolar depressed subjects (a diagnosis which is also prominent in the translocation family). Such a study provides an opportunity to investigate a possible association between the translocation breakpoint region and a postulated schizophrenia susceptibility gene in a random population of affected subjects versus normal controls. No association was demonstrated in either the schizophrenia subjects or the unipolar depressed subjects, evidence against a nearby gene of major effect. As alluded to in the discussion section of chapter 5, this may imply that either the postulated gene in close proximity to the breakpoint is not in linkage disequilibrium with the marker used or that the gene is of major effect in the translocation family, but not in the general population of schizophrenics tested. Many of the issues related to the procedural aspects of association studies have been previously referred to, such as that of population stratification (see chapter 5 discussion). However, the question remains of what should be regarded as positive/negative association in terms of statistical significance and, if present, what this difference means (Berrettini 1997, Paterson 1997, Baron 1997). These issues have been widely discussed in the literature, particularly that of statistical significance (Lander and Kruglyak 1995, Carey 1994).

Since we lack any clear evidence to distinguish candidate genes we must rely heavily on statistical argument and on what is technically and practically possible. The level of statistical significance is ideally set at a value which does not preclude the detection of true small associations but that does eliminate the vast majority of false negative results. However, this precise value at which it should be set is highly debated in the literature. If a p value is set at 0.05 then by definition the observed finding would have occurred by chance one time in twenty if the null hypothesis were true. False positive results are, according to Crowe (1993) more likely to occur due to the low prior probability of any chosen gene being involved in neuropsychiatric disorder (due to our ignorance of the complex genetic architecture of the higher central nervous system). For each candidate gene considered there

are probably hundreds of equally likely genes which we, as yet, have not identified. Crowe demonstrates that because the prior likelihood is so low, the significance level required to have a 5% false positive rate equates to a p value of <0.00001. Further statistical complications arise from the concept of multiple testing (the more loci tested the more likely that a false positive result will be produced purely by chance). This is a serious complication at a time when it is practical to carry out high density genome scans (multiple testing) but there is little to distinguish one or other candidate gene from the long list of possible candidates (specific locus testing). Such complications have lead many researchers to conclude that the only way to interpret positive association findings is to consistently replicate the result in several populations and thus confirm the "true" association (Paterson 1997, Baron 1997).

However, if further studies in ethnically diverse populations fail to replicate the original positive findings, this does not negate the original result, since genetic heterogeneity and several different biochemical mechanisms and pathways may be responsible and indeed is highly likely in psychiatric diseases. The degree of locus heterogeneity and the number of genes involved in many complex diseases may be high. If in different populations different genes are involved, then this would result in unreplicable positive results from different samples. Similarly, allelic variants which are not associated in one population may be associated with the disease in a different population and therefore for the same reasons negative results can not be ruled out.

The reported association between alcoholism and the TaqI A1 RFLP in the dopamine D2 receptor gene (Blum et al 1990), provides a pertinent example of the difficulties that can be incurred in the replication of association studies and highlights many of the problems in the design of association studies. Several groups have replicated the original positive association (Blum et al 1990, Comings et al 1991, Amadeo et al 1993) but many others have failed to do so (Turner et al 1992, Arinami et al 1993, Suarez et al 1994). One suggestion for the failure of some groups to replicate the original finding is due to the differences in phenotype definition which varies slightly from study to study. Exact phenotype definition is problematic in psychiatric disorders such as schizophrenia and any replication attempts of positive association studies should follow closely the definition used in the original study.

Further explanations for failure to demonstrate consistent replication have included;

a) population stratification

b) the use of "Super normal" controls (which have been screened to exclude all psychopathology and may result in over-estimation of comorbid association)

c) a related phenotype, which as yet has not been defined, is associated to this marker (Neiswanger et al 1995).

Many of these problems can be overcome by using family based association methods. Interestingly when these methods have been employed in the case of the TaqI A1 allele and alcoholism, the results have been negative as have those of linkage analysis, lending support to the artefactual nature of the initial positive associations. (It is important to note that association and linkage may not always occur together for several reasons including the fact that linkage is more likely to determine genes of moderate to major effect). However, population based association studies are relatively easy and inexpensive to carry out and once collected, many markers can be typed with relative ease. In this respect such population based studies are still appealing. Since association studies are so dependent on statistical analysis by their very nature, large sample populations should be employed to improve the power of such analysis. There is also undoubtedly a role for pooling samples and meta-analysis of data produced.

The fourth and final approach described in this thesis was to directly isolate gene transcripts around the region of the chromosome 1 breakpoint employing a cDNA selection technique, coincident sequence cloning. By using foetal brain cDNA and cloned genomic DNA from the region of the chromosome 1 breakpoint, in the form of a PAC and 2 cosmids, it was possible to isolate gene fragments expressed in foetal brain from this region of the breakpoint. This technique allowed the isolation of two fragments which will require further investigation to determine their candidacy. Both of these fragments lie in relatively close proximity to the chromosome 1 breakpoint. One appears to be novel, as judged by database searching, the other shows strong identity to two ESTs in the database. One of these ESTs is expressed in foetal lung and the other in foetal heart. Future work resulting from this thesis will primarily consist of the continued investigation of these two clones as alluded to in the discussion of chapter 6.

The four approaches taken in this thesis have proved successful in including or excluding gene fragments residing in proximity to the breakpoint from further study. Exclusion of fragments is of course as important as inclusion, particularly in regions in which there are a plethora of genes and no strong prior hypotheses regarding function. The majority of the gene fragments considered in this thesis have been excluded based on their remote location from the breakpoint rather than on their biological function. With the exception of a gene fragment which spans the breakpoint and is disrupted by it, it will of course become increasingly difficult to exclude gene fragments which lie in close proximity to the breakpoint particularly when they are of unknown function. Exclusion of such gene fragments will require additional work on more functional aspects of these genes such as further analysis of the DNA sequence and translated product sequence (to determine functional domains) as well as assessment of expression patterns (RT-PCR and RNA *insitu*).

The examples of the difficulties incurred when using these approaches indicates the importance of using differing strategies to solving difficult problems. Different approaches allow the independent confirmation of a result which is increasingly important in the study of complex disorders.

Future work resulting from this thesis will consist primarily of following up the clones ascertained from the CSC experiment (see chapter 6 discussion).

7.2 Progress in the physical mapping of the t(1:11) translocation breakpoint and flanking genomic regions during the course of this thesis:

During the course of this thesis, work carried out by other members of the psychiatric genetics group has continued and advanced the physical mapping of the translocation breakpoint region.

On chromosome 11, once the breakpoint spanning YAC D0485 had been identified it was screened for genes using a coincident sequence cloning strategy. This resulted in the identification of several gene fragments. These included a tubulin pseudogene which mapped ~250Kb proximal to the chromosome 11 breakpoint as well as a number of novel fragments. These fragments have been further characterised and extended in sequence. The novel fragments made up three novel genes which resided >500Kb proximal to the chromosome 11 translocation breakpoint.

The YAC D0485 was also subcloned into cosmids. The cosmid sub-clones were used to produce a contig around the chromosome 11 breakpoint region. Three cosmids were identified as crossing the breakpoint by FISH on metaphase chromosomes from a cell line bearing der 1 and der 11 chromosomes (MAFLI). Restriction enzyme mapping indicated a limited overlap between these cosmids. *EcoRI* restriction enzyme digestion identified three common fragments from these three cosmids. The largest fragment (2.5Kb) contained a repeat element which was removed by *HindIII* digestion, producing a 2.15Kb fragment. This fragment was shown to span the translocation breakpoint by FISH on the somatic cell hybrid MAFLI, and by Southern blot hybridisation to a somatic cell hybrid panel containing the der1 and der11 only hybrids MIS39.8 and MIS7.4. This 2.15Kb fragment hybridised to a single 2.5Kb fragment in human genomic control DNA but to two additional fragments at 2.7Kb (resulting from the der1 chromosome) and 7Kb (from der11 chromosome) in DNA derived from an affected family member carrying the translocation chromosomes.

In collaboration with a past member of the Edinburgh group (Benoit Arvileiler, Bordeaux) a chromosome 1 YAC contig was established using several of the microdissection clones which mapped to chromosome 1 to isolate YACs from commercial YAC libraries. Marker analysis using the microdissection clones, YAC end clones and published AFM markers allowed the contig to be attained.

In order to cross over to chromosome1, the 2.15Kb wild-type 11 fragment was used as a probe to screen a genomic DNA library made from a family member carrying the translocation chromosomes (MAFLI). This resulted in the isolation of the der 1 breakpoint spanning clone. This clone was then used as a probe onto the library and resulted in the isolation of the chromosome 1 wild-type breakpoint fragment (7.3Kb). The der 11 breakpoint fragment was also isolated from this library using the wild-type 11 fragment. The sequence of the breakpoint spanning clones showed that the breakpoint did not contain any substantial deletions or insertions. However, above the breakpoint on chromosome 11, the sequence TCAG was substituted with AA. Otherwise, the breakpoint appeared to be clean. No mechanism for the occurrence of the breakpoint event (such as repeat homology on either side of the breakpoint) was apparent from the sequence obtained around the breakpoint on either chromosome 1 or 11. The wild-type chromosome 1 7.3 Kb sequence

indicated several exon predictions as determined by GRAIL (Xu et al 1996) and several EST matches were detected when the sequence was compared to sequences in the EST and GENBANK databases. One of the exons has been shown to cross the breakpoint on chromosome 1. As such, this cDNA fragment constitutes the best candidate gene for involvement in the psychiatric diagnosis in the translocation family and is currently being extended to full length sequence and is undergoing expression analysis by RT-PCR and RNA *insitu*.

The isolation of a gene fragment which spans the breakpoint on chromosome 1 provides an excellent candidate gene for involvement in the psychiatric diagnosis in the translocation family. Characterisation of this gene and its associated elements will hopefully reveal a great deal about the effect the translocation event has on the function of this gene and in turn elucidate the role this gene may play in evoking the psychiatric diagnosis seen in the translocation family.

The chromosome 1 YAC contiguous clone map has been further added to by screening cosmid and PAC libraries from chromosome 1 with probes identified from the chromosome 1 breakpoint region. This resulted in the identification of a PAC and several cosmids which reside around the chromosome 1 breakpoint region. Further information on genes located in the breakpoint region was obtained from the Human Transcript Map (Schuler et al 1996) which indicated that several ESTs resided in the general region around the chromosome 1 breakpoint (as determined by STS markers mapped to the region).

The number of ESTs mapped to the general region of the chromosome 1 YAC contig was large (>40) and seem to indicate that this region of chromosome 1 is extremely gene rich in comparison with chromosome 11 which appears to have a paucity of genes. (Of course many of these ESTs may represent parts of the same gene). However, the ESTs from the Human Transcript Map were located at distances >40Kb from the breakpoint. To date the cDNAs in close proximity to the chromosome 1 breakpoint and the cDNA which has been shown to span the breakpoint, are the highest priority candidates for involvement in the psychiatric diagnosis in the translocation family and, as such, will be vigorously pursued to elucidate their role in this process.

7.3 How close are we to finding a schizophrenia gene?

Schizophrenia has to date proved refractory to all attempts to extract its underlying genetic cause. Non Mendelian inheritance, genetic heterogeneity, uncertainty of phenotype and incomplete penetrance all conspire to make schizophrenia a complex and challenging disorder to investigate. Until recently, the progress made in investigating the genetics of this disorder were disillusioning with advances being followed closely by retreat. Recently, changes in methodology and inventive new modifications of older method designs have cracked opened the door to the mystery of schizophrenia, if only slightly. Replicated data from large collaborative linkage studies (as well as smaller studies) have indicated regions on chromosome 6p (Straub et al 1995, Wang et al 1995, Moises et al 1995, Antonarakis et al 1995), 8p (Pulver et al 1995, Moises et al 1995) and 22q (Pulver et al 1994, Coon et al 1994, Vallda et al 1995) as possible regions for the presence of a susceptibility gene for schizophrenia. Chromosome 22q12-q13 has also demonstrated statistically significant association with markers D22S283 and D22S278 lending further support to the possible location of a schizophrenia susceptibility gene in this region. However, every positive linkage and association result has been accompanied by a negative result and although encouraging, more work is needed to confirm these results.

Further interesting data has come from work on the dopamine D3 receptor gene polymorphism in exon 1. A number of groups have reported excess homozygosity in the patient group compared to that of the control group (Crocq et al 1992, Mant et al 1994, Nimgaonkar et al 1993). However, the familiar story of several other groups failing to find this observation (Jonsson et al 1993, Yang et al 1993, Nothen et al 1993) is also apparent in this case. The T to C polymorphism at nucleotide 102 of the serotonin receptor with schizophrenia reported in a Japanese study (Inayama et al 1994, Williams et al 1996) is also of interest, not least because (as with the dopamine D3 receptor polymorphism) there are many unanswered questions about the relationship of the polymorphism with the neuroreceptor genes and schizophrenia. Further work on the functional nature of these polymorphisms is required to establish the nature of this relationship.

The recent interest in expanded trinucleotide repeats (the hallmark of which is the phenomenon of anticipation) has added yet another thread to the web of intrigue

(Gorwood et al 1996, O'Donnovan et al 1996, Vincent et al 1996). Although the phenomena of anticipation in schizophrenia is widely debated, as many believe that systematic sampling biases are responsible for anticipation being reported, several groups have provided molecular data that some trinucleotide repeats (CAG/CTG) are expanded in schizophrenia subjects (O'Donovan et al 1995, Morris et al 1995). The size of contribution to the overall variance of schizophrenia is however unclear but as with all the other examples mentioned above is likely to be relatively small.

We will doubtless learn much from the rapidly progressing research being carried out in neurodegenerative disorders such as Alzheimer's disease and Huntington's disease. Elucidation of the pathological mechanisms by which such specific neurone damage can occur in these disorders will advance our knowledge of the pathophysiological processes involved in such late onset neurodegenerative disorders. Both of these disorders, particularly Huntingtons disease, have indicated new mechanisms (i.e. triplet repeat expansions) by which brain function can be specifically altered. Increasing the general knowledge of the ways in which such alterations can occur will allow improved understanding of the complex mechanisms which may interact to result in altered function of the brain.

Positional cloning of the t(1;11)translocation breakpoint in the K26 pedigree provides an alternative approach for determining schizophrenia susceptibility genes, to the more traditional linkage analysis approaches being employed in the field. This study allows the hypothesis that the translocation event in some way disrupts a gene in close proximity to the breakpoint resulting in the diagnosis of mental illness in this family, to be directly tested. By isolating genes in the region of the breakpoint it may be possible to determine how the pathological process which leads to the psychiatric diagnosis in this family occurs. Description of a gene which is causally associated with mental illness, albeit in this one family, would represent a significant step towards the ultimate aim of understanding the multiple interacting genetic and environmental effects which lead to schizophrenia.

Future research into the complex pathophysiology of schizophrenia may benefit from complementary strategies currently being employed in the field which are converging to produce interesting data on the brain abnormalities which underlie the disease. Such research includes the plethora of work being produced from

functional brain imaging studies (reviewed by Velakoulis and Pantelis 1996). Such studies allow visualisation of brain areas that are activated by a variety of sensory, motor or cognitive tasks. Metabolic processes linked to neuronal activity such as cerebral blood flow, cerebral metabolism as well as receptor occupancy and density etc. are most often looked at in such studies. Functional imaging technology is providing increasing insights into schizophrenia and its treatment, proving to be a reliable, reproducible non invasive technique for visualising regions of the brain which may be involved in psychiatric disorders.

Studies which dissect complex human disease phenotypes into constituent behaviours in animal models are currently finding application in psychiatric genetics. These approaches utilise gene targeting and selectively bred recombinant inbred and transgenic strains and are a powerful way of mapping genes of lesser effect which influence behaviour. An example of such an approach is the recent work by Lijam et al (1997) in which mice completely deficient for *Dvl1* (one of three mouse homologs of the Drosophila segment polarity gene *Deshevelled)* were created. These mice were shown to exhibit reduced social interaction and had sensorimotor gating abnormalities. Such abnormalities have been seen in patients with schizophrenia and other mental disorders and therefore *Dvl1* mouse mutants may provide a model for aspects of several human psychiatric disorders. Ultimately it is hoped that this may lead to the identification of candidate loci for genetic analysis in humans and provide a model system to study putative disease pathophysiology.
Transgenic animal approaches are also helping to clarify the actions of various candidate genes such as the dopamine D4 receptor (DRD4) which has previously received considerable attention due to its high affinity for the atypical antipsychotic clozapine (Seeman and Van Tol 1994) and the polymorphic nature of this gene (Van Tol et al 1992). No allele of the DRD4 receptor has been associated with any particular mental disorder but there have been a number of reports showing that certain alleles may predispose an individual to alcoholism (Muramatsu et al 1996), novelty seeking behaviour (Ebstein et al 1996) and opiate abuse (Kotler et al 1997). Mice which lack this protein ($DRD4^{-/-}$) have been produced to clarify the in vivo role of the DRD4 receptor (Rubinstein et al 1997). The results from this study Rubinstein et al 1997) demonstrated that the mutant DRD4 deficient mice were supersensitive to ethanol, cocaine and methamphetamine and less active that wildtype mice in open field tests. Biochemical analysis showed that dopamine synthesis and its

conversion to DOPAC were elevated in the dorsal stratium compared to wildtype mice. The authors propose that DRD4 acts at the cellular level as an inhibitory postsynaptic receptor that primarily modulates the firing of neurones in the frontal cortex and basal ganglia. DRD4 therefore could modulate normal, co-ordinated and drug-stimulated motor behaviours as well as the activity of nigrostriatal dopamine neurons. However, more studies are needed to confirm this.

A recent interesting animal study was carried out by Jentsch et al (1997) in which they treated vervet monkeys with phenylcyclidine (PCP) twice a day for fourteen days. PCP can cause schizophrenia symptoms in humans and if used repeatedly can result in a long lasting syndrome marked by neuropsychological deficits, social withdrawal and affective blunting. Cognitive dysfunction (seen in schizophrenia) and frontal lobe damage is also observed in long term abuse of PCP. Since this may represent a pharmacological model of cognitive deficits of the frontal cortex that are associated with schizophrenia, the behavioural and neurochemical effects of long term exposure to PCP were investigated by this group (Jentsch et al 1997). They reported that monkeys treated with PCP for fourteen days displayed performance deficits on tasks that were sensitive to prefrontal cortex function and that these deficits were ameliorated by the atypical antipsychotic drug clozapine. Biochemical studies showed that repeated exposure to PCP caused a reduction in both basal and evoked dopamine utilisation in the dorsolateral prefrontal cortex (associated with cognitive function). These effects were demonstrated not to be the result of direct drug effects, continuing after administration of PCP was stopped. This model may therefore prove useful to study psychiatric disorders, such as schizophrenia, which are associated with cognitive dysfunction and dopamine hypofunction in the prefrontal cortex.

Several areas of research are beginning to converge to provide important information about several aspects of psychiatric illness, demonstrating that taking a multidisciplinary approach to solving complex problems is beneficial. Convergence of all these areas of science - biochemistry, pharmacology, functional imaging analysis, genetics and many others, provides a holistic approach to solving the mystery of schizophrenia which will benefit those afflicted with the disorder. Greater understanding of the pathophysiology of this disorder will lead to improved, more effective treatments for the majority of sufferers. Such progress is both

economically and socially desirable and will hopefully lead to a better quality of life for sufferers and their families who are at present left to cope with an inadequate system of "care in the community". With greater understanding of the condition and its cause perhaps the community will begin to care and patients and their families will receive the support and understanding that this condition deserves.

References

References

Abrams, R. and Taylor, M.A. (1983): The genetics of schizophrenia: a reassessment using modern criteria. *Am.J.Psychiatry*, 140:171-175.

Altschul, S.F. and Gish, W. (1990): Basic local alignment search tool. *J.Mol.Biol.*, 215:403-410.

Amadeo, S., Abbar, M., Fourcade, M.L., Waksman, G., Leroux, M.G., Madec, A., Selin, M., Champiat, J.C., Brethome, A., Leclaire, Y., Castelnau, D., Venisse, J.L., and Mallet, J. (1993): D2 dopamine-receptor gene and alcoholism. *Journal Of Psychiatric Research*, 27:173-179.

American Psychiatric Association (1980) *Diagnostic and statistical manual of mental disorders.* 3rd Edition Washington D.C.

American Psychiatric Association (1987): *Diagnostic and statistical manual of mental disorders.* 3rd Edition Rev. Washington D.C.

American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders*, 4th edition, American Psychiatric Association, Washington D.C

Anand, R., Riley, J.H., Butler, R., Smith, J.C., and Markham, A.F. (1990): A 3.5 genome equivalent multi access YAC library: construction, characterisation and storage. *Nucl.Acids.Res.*, 19:1951-1956.

Andreasen, N.C. (1988): Brain imaging - applications in psychiatry. *Science*, 239:1381-1388.

Antonarakis, S.E., Blouin, J.L., Pulver, A.E., Wolyniec, P., Lasseter, V.K., Nestadt, G., Kasch, L., Babb, R., Kazazian, H.H., Dombroski, B., Kimberland, M., Ott, J., Housman, D., Karayiorgou, M., and Maclean, C.J. (1995): Schizophrenia

susceptibility and chromosome 6p24-22. *Nature Genetics*, 11:235-236.

Antonarakis, S.E., Blouin, J.L., Pulver, A.E., Wolyniec, P., Lasseter, V.K., Nestadt, G., Babb, R., Kazazian, H.H., Dombroski, B., Kimberland, M., Ott, J., Karayiorgou, M., Housman, D., and Maclean, C. (1995): A potential susceptibility locus for schizophrenia on chromosome 6p24-p22. *American Journal Of Human Genetics*, 57:1061-1061.

Arinami, T., Itokawa, M., Komiyama, T., Mitsushio, H., Mori, H., Mifune, H., Hamaguchi, H., and Toru, M. (1993): Association between severity of alcoholism and the a1-allele of the dopamine d2-receptor gene taqi-a rflp in japanese. *Biological Psychiatry*, 33:108-114.

Arnold, S.E., Lee, V.Y., Gur, R.E., and Trojanowski, J.Q. (1991): Abnormal expression of 2 microtubule-associated proteins (Map2 and map5) In specific subfields of the hippocampal-formation in schizophrenia. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 88:10850-10854.

Arveiler, B. and Porteous, D. (1992): Distribution of Alu and L1 repeats in human YAC recombinants. *Mammalian Genome*, 3:661-668.

Asherson, P., Parfitt, E., Sargeant, M., Tidmarsh, S., Buckland, P., Taylor, C., Clements, A., Gill, M., McGuffin, P., and Owen, M. (1992): No evidence for a pseudoautosomal locus for schizophrenia. Linkage analysis of multiply affected families. *Br.J.Psychiatry*, 161:63-68.

Babitch, J.A. and Zheng, X. (1995): Subcellular-distribution of alpha-actinin in chick brain. *Journal Of Neurochemistry*, 65:S-S

Baileywilson, J.E. and Bamba, V. (1993): Sib-pair linkage analyses of alzheimers-disease. *Genetic Epidemiology*, 10:371-376.

Barnes, D.M. (1986): Brain architecture:beyond genes. *Nature.*, 233:155-156.

Baron, M. (1976): Albinism and schizophreniform psychosis: a pedigree study. *Am.J.Psychiat.*, 133:1070-1073.

Baron, M., Gruen, R., Rainer, J.D., Kane, J., Asnis, L., and Lord, S. (1985): A family study of schizophrenic and normal control probands: implications for the spectrum concept of schizophrenia. *Am.J.Psychiat.*, 142:447-455.

Baron, M. and Gruen, R.S. (1991): Schizophrenia and affective disorder: are they genetically linked? *Br.J.Psychiatry*, 159:267-270.

Baron, M. (1997): Association studies in psychiatry: A season of discontent. *Molecular Psychiatry*, 2:278-281.

Baron, M.D., Davison, M.D., Jones, P., Patel, B., and Critchley, D.R. (1987): Isolation and characterization of a cdna-encoding a chick alpha- actinin. *Journal of Biological Chemistry*, 262:2558-2561.

Baron, M.D., Davison, M.D., Jones, P., and Critchley, D.R. (1987): The structure and function of alpha-actinin. *Biochemical Society Transactions*, 15:796-798.

Baron, M.D., Davison, M.D., Jones, P., and Critchley, D.R. (1987): The sequence of chick alpha-actinin reveals homologies to spectrin and calmodulin. *Journal of Biological Chemistry*, 262:17623-17629.

Barr, C.L., Kennedy, J.L., Pakstis, A.J., Wetterberg, L., Sjogren, B., Gelernter, J., Hallmayer, J., Moises, H., Cavallisforza, L.L., and Kidd, K.K. (1991): Progress in genome scan for linkage in schizophrenia. *American Journal Of Human Genetics*, 49:335-335.

Barr, C.L., Kennedy, J.L., Lichter, J.B., Van Tol, H.H., Wetterberg, L., Livak, K.J., and Kidd, K.K. (1993): Alleles at the dopamine D4 receptor locus do not

contribute to the genetic susceptibility to schizophrenia in a large Swedish kindred. *Am.J.Med.Genet.*, 48:218-222.

Barr, C.L., Kennedy, J.L., Pakstis, A.J., Wetterberg, L., Sjogren, B., Bierut, L., Wadelius, C., Wahlstrom, J., Martinsson, T., Giuffra, L., Gelernter, J., Hallmayer, J., Moises, H.W., Kurth, J., Cavalli-Sforza, L.L., and Kidd, K.K. (1994): Progress in a genome scan for linkage in schizophrenia in a large Swedish kindred. *Am.J.Med.Genet.*, 54:51-58.

Bassett, A.S., McGillivary, B.C., Jones, B.D., and Pantzar, J.T. (1988): Partial trisomy chromosome 5 cosegregating with schizophrenia. *Lancet*, 1:799-801.

Bassett, A.S. and Honer, W.G. (1994): Evidence for anticipation in schizophrenia. *Am.J.Hum.Genet.*, 54:864-870.

Bassett, A.S., Bury, A., Hodgkinson, K.A., and Honer, W.G. (1996): Reproductive fitness in familial schizophrenia. *Schizophrenia Research*, 21:151-160.

Bastard, C., Deweindt, C., Kerckaert, J.P., and Tilly, H. (1993): Laz3, the gene involved in 3q27 translocations is rearranged in about 30-percent of diffuse b-cell lymphoma. *Blood*, 82:A 440-A 440

Battaglia, M. and Torgersen, S. (1996): Schizotypal disorder - at the crossroads of genetics and nosology. *Acta Psychiatrica Scandinavica*, 94:303-310.

Bebbington, P. and Kuipers, L. (1994): The clinical utility of expressed emotion in schizophrenia. *Acta Psychiatr.Scand.Suppl.*, 382:46-53.

Beggs, A.H., Hoffman, E.P., Snyder, J.R., Arahata, K., Specht, L., Shapiro, F., Angelini, C., Sugita, H., and Kunkel, L.M. (1991): Exploring the molecular basis for variability among patients with Becker muscular dystrophy: dystrophin gene and protein studies. *Am.J.Hum.Genet.*, 49:54-67.

Beggs, A.H., Byers, T.J., Knoll, J.H.M., Boyce, F.M., Bruns, G.A.P., and Kunkel, L.M. (1992): Cloning and characterisation of two human skeletal muscle alpha-actinin genes located on chromosomes one and eleven. *J.Biol.Chem.*, 267:9281-9288.

Benjamin, J., Li, L., Patterson, C., Greenberg, B.D., Murphy, D.L., and Hamer, D.H. (1996): Population and familial association between the d4 dopamine-receptor gene and measures of novelty seeking. *Nature Genetics*, 12:81-84.

Bennett, V., Davis, J., and Fowler, W.E. (1982): Brain spectrin, a membrane-associated protein related in structure and function to erythrocyte spectrin. *Nature*, 299:126-131.

Beratis, S., Gabriel, J., and Hoidas, S. (1994): Age at onset in subtypes of schizophrenic disorders. *Schizophr.Bull.*, 20:287-296.

Berrettini, W. (1997): On the interpretation of association studies in behavioral disorders. *Molecular Psychiatry*, 2:274-275.

Bird, A.P. (1986): CpG-rich islands and the function of DNA methylation. *Nature.*, 321:209-213.

Blackwood, D.H.R., St.Clair, D.M., Muir, W.J., and Duffy, J.C. (1991): Auditory P300 and eye tracking dysfunction in schizophrenic pedigrees. *Arch.Gen.Psychiatry.*, 48:899-909.

Blanchard, A., Ohanian, V., and Critchley, D. (1989): The structure and function of alpha-actinin. *Journal Of Muscle Research And Cell Motility*, 10:280-289.

Bleich, A., Brown, S.L., Kahn, R., and Vanpraag, H.M. (1988): The role of serotonin in schizophrenia. *Schizophrenia Bulletin*, 14:297-315.

Blennow, G. and McNeil, T.F. (1991): Neurological deviations in newborns at psychiatric high risk. *Acta Psychiatr.Scand.*, 84:179-184.

Bleuler E Dementia praecox or group of schizophrenias (1911) Trans joseph Zinkin New York International Universities 1950

Boguski, M.S. and Schuler, G.D. (1995): ESTablishing a human transcript map. *Nat.Genet.*, 10:369-371.

Bracha, H.S., Torrey, E.F., Bigelow, L.B., Lohr, J.B., and Linington, B.B. (1991): Subtle signs of prenatal maldevelopment of the hand ectoderm in schizophrenia: a preliminary monozygotic twin study. *Biol.Psychiatry*, 30:719-725.

Bradbury, T.N. and Miller, G.A. (1985): Season of birth in schizophrenia - a review of evidence, methodology, and etiology. *Psychological Bulletin*, 98:569-594.

Brookes, A.J., Slorach, E.M., Morrison, K.E., Qureshi, S.J., Blake, D., Davies, K., and Porteous, D.J. (1994): Cloning the shared components of complex DNA resources. *Hum.Mol.Gen.*, 3:2011-2017.

Brown, W.A. and Bird, A.P. (1986): Long-range restriction site mapping of mammalian genomic dna. *Nature*, 322:477-481.

Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., and Sharp, P.A.H.D.E. (1991): Exon amplification; A strategy to isolate mammalian genes based on RNA splicing. *Proc.Natl.Acad.Sci.USA*, 88:4005-4009.

Burton, F.H., Loeb, D.D., Voliva, C.F., Martin, S.L., Edgell, M.H., and Hutchison, C.A.3d. (1986): Conservation throughout mammalia and extensive protein-encoding capacity of the highly repeated DNA long interspersed sequence one. *J.Mol.Biol.*, 187:291-304.

Cambien, F., Poirier, O., Lecerf, L., Evans, A., Cambou, J.P., Arveiler, D., Luc, G., Bard, J.M., Bara, L., Ricard, S., Tiret, L., Amouyel, P., Alhencgelas, F., and Soubrier, F. (1992): Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial-infarction. *Nature*, 359:641-644.

Cambien, F. and Soubrier, F. (1992): The insertion/deletion polymorphism of the angiotensin-i converting enzyme appears to be a risk factor of myocardial-infarction. *M S-Medecine Sciences*, 8:989-991.

Campion, D., d'Amato, T., Bastard, C., Laurent, C., Guedj, F., Jay, M., Dollfus, S., Thibaut, F., Petit, M., Gorwood, P., Babron, M.C., Waksman, G., Martinez, M., and Mallet, J. (1994): Genetic study of dopamine D1, D2, and D4 receptors in schizophrenia. *Psychiatry Res.*, 51:215-230.

Cao, Q.H., Martinez, M., Zhang, J., Sanders, A.R., Badner, J.A., Cravchik, A., Markey, C.J., Beshah, E., Guroff, J.J., Maxwell, M.E., Kazuba, D.M., Whiten, R., Goldin, L.R., Gershon, E.S., and Gejman, P.V. (1997): Suggestive evidence for a schizophrenia susceptibility locus on chromosome 6q and a confirmation in an independent series of pedigrees. *Genomics*, 43:1-8.

Carey, G. (1994): Genetic association study in psychiatry: analytical evaluation and a recommendation. *Am.J.Med.Genet.*, 54:311-317.

Carter, M. and Watts, C.A.H. (1971): Possible biological advantages among schizophrenics' relatives. *Br.J.Psychiatry*, 118:453-460.

Castle, D.J. and Murray, R.M. (1991): The neurodevelopmental basis of sex differences in schizophrenia. *Psychol.Med.*, 21:565-575.

Chen, C.H., Lee, Y.R., Wei, F.C., Koong, F.J., Hwu, H.G., and Hsiao, K.J. (1997): Lack of allelic association between 102T/C polymorphism of serotonin receptor

type 2A gene and schizophrenia in Chinese. *Psychiatric Genetics*, 7:35-38.

Chen, Z., Brand, N.J., Chen, A., Chen, S.J., Tong, J.H., Wang, Z.Y., Waxman, S., and Zelent, A. (1993): Fusion between a novel kruppel-like zinc finger gene and the retinoic acid receptor-alpha locus due to a variant t(11,17) Translocation associated with acute promyelocytic leukemia. *Embo Journal*, 12:1161-1167.

Chou, Q., Russell, M., Birch, D.E., Raymond, J., and Bloch, W. (1992): Prevention of PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucl.Acids Res.*, 20:1717-1723.

Clarke, D.J. and Buckley, M.E. (1989): Familial association of albinism and schizophrenia. *Br.J.Psychiatry*, 155:551-553.

Cleghorn, J.M., Zipursky, R.B., and List, S.J. (1991): Structural and functional brain imaging in schizophrenia. *J.Psychiatry Neurosci.*, 16:53-74.

Cohen, J. (1988): *Statistical power analysis of the behaioural sciences*. Erlbaum Associates, Hillsdale, New Jersey.

Collinge, J., DeLisi, L.E., Boccio, A., Johnstone, E.C., Lane, A., Larkin, C., Leach, M., Lofthouse, R., Owen, F., Poulter, M., Shah, T., Walsh, C., and Crow, T.J. (1991): Evidence for a pseudo-autosomal locus for schizophrenia using the method of affected sib pairs. *Br.J.Psychiatry*, 158:624-629.

Collins, F.S. (1995): Positional cloning moves from perditional to traditional. *Nature Genetics*, 9:347-350.

Comings, D.E., Comings, B.G., Muhleman, D., Dietz, G., Shahbahrami, B., Tast, D., Knell, E., Kocsis, P., Baumgarten, R., Kovacs, B.W., Levy, D.L., Smith, M., Borison, R.L., Evans, D.D., Klein, D.N., Macmurray, J., Tosk, J.M., Sverd, J., Gysin, R., and Flanagan, S.D. (1991): The dopamine-d2 receptor locus as a modifying gene in neuropsychiatric disorders. *Jama-Journal Of The American*

Medical Association, 266:1793-1800.

Coon, H., Hoff, M., Holik, J., DeLisi, L.E., Crowe, T., Freedman, R., Shields, G., Boccio, A.M., Lerman, M., and Gershon, E.S. (1993): C to T nucleotide substitution in codon 713 of amyloid precursor protein gene not found in 86 unrelated schizophrenics from multiplex families. *Am.J.Med.Genet.*, 48:36-39.

Coon, H., Sobell, J., Heston, L., Sommer, S., Hoff, M., Holik, J., Umar, F., Robertson, M., Reimherr, F., Wender, P., Vest, K., Myles-Worsley, M., Gershon, E.S., DeLisi, L.E., Shields, G., Dale, P.W., Polloi, A., Waldo, M., Leonard, S., Sikela, J., Freedman, R., and Byerley, W. (1994): Search for mutations in the beta 1 GABAA receptor subunit gene in patients with schizophrenia. *Am.J.Med.Genet.*, 54:12-20.

Corder, E.H., Saunders, A.M., Risch, N.J., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Rimmler, J.B., Locke, P.A., Conneally, P.M., Schmader, K.E., Small, G.W., Roses, A.D., Haines, J.L., and Pericakvance, M.A. (1994): Apolipoprotein-e type-2 allele decreases the risk of late-onset Alzheimer-disease. *Neurobiology Of Aging*, 15:S-S

Crocq, M.A., Mant, R., Asherson, P., Williams, J., Hode, Y., Mayerova, A., Collier, D., Lannfelt, L., Sokoloff, P., Gill, M., Macher, J.P., McGuffin, P., and Owen, M.J. (1992): Association between schizophrenia and homozygosity at the dopamine D3 receptor gene. *J.Med.Genet.*, 29:858-860.

Crow, T.J. (1980): Molecular pathology of schizophrenia: more than one disease process. *British Medical Journal*, 280:66-68.

Crow, T.J. (1988): Sex-chromosomes and psychosis - the case for a pseudoautosomal locus. *British Journal of Psychiatry*, 153:675-683.

Crow, T.J. (1990): Temporal lobe assymetries as the key to the etiology of schizophrenia. *Schiz.Bull.*, 16(3):433-443.

Crowe, R.R., Black, D.W., Wesner, R., Andreasen, N.C., Cookman, A., and Roby, J. (1991): Lack of linkage to chromosome 5q11-q13 markers in six schizophrenia pedigrees. *Arch.Gen.Psychiatry*, 48:357-361.

Crowe, R.R. (1993): Candidate genes in psychiatry - an epidemiologic perspective. *American Journal Of Medical Genetics*, 48:74-77.

Csernansky, J.G., Murphy, G.M., and Faustman, W.O. (1991): Limbic mesolimbic connections and the pathogenesis of schizophrenia. *Biological Psychiatry*, 30:383-400.

Cullberg, J. (1991): Recovered versus nonrecovered schizophrenic-patients among those who have had intensive psychotherapy. *Acta Psychiatrica Scandinavica*, 84:242-245.

d'Amato, T., Campion, D., Gorwood, P., Jay, M., Sabate, O., Petit, C., Abbar, M., Malafosse, A., Leboyer, M., Hillaire, D., Clerget-Darpoux, F., Feingold, J., Waksman, G., and Mallet, J. (1992): Evidence for a pseudoautosomal locus for schizophrenia. II: Replication of a non-random segregation of alleles at the DXYS14 locus. *Br.J.Psychiatry*, 161:59-62.

Dalby, J.T., Morgan, D., and Lee, M.L. (1986): Schizophrenia and mania in identical twin brothers. *THe Journal of Nervous and Mental Disease*, 174:304-308.

Dallery, E., Galieguezouitina, S., Collyndhooghe, M., Quief, S., Denis, C., Hildebrand, M.P., Lantoine, D., Deweindt, C., Tilly, H., Bastard, C., and Kerckaert, J.P. (1995): Ttf, a gene encoding a novel small g-protein, fuses to the lymphoma- associated laz3 gene by t(3-4) chromosomal translocation. *Oncogene*, 10:2171-2178.

Dassa, D., Sham, P.C., Vanos, J., Abel, K., Jones, P., and Murray, R.M. (1996): Relationship of birth season to clinical-features, family history, and obstetric complications in schizophrenia. *Psychiatry Research*, 64:11-17.

Dassa, D., Azorin, J.M., Ledoray, V., Sambuc, R., and Giudicelli, S. (1996): Season of birth and schizophrenia - sex difference. *Progress In Neuro-Psychopharmacology & Biological Psychiatry*, 20:243-251.

Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, L.E., Reed, P.W., Gough, S.L., Jenkins, S.C., Palmer, S.M., Balfour, K.M., Rowe, B.R., Farrall, M., Barnett, A.H., Bain, S.C., and Todd, J.A. (1994): A genome-wide search for human type-1 diabetes susceptibility genes. *Nature*, 371:130-136.

Davis, K.L. (1991): Dopamine in schizophrenia: a review and reconceptualization. *Am.J.Psychiatry*, 148:1474-1486.

Deihl ,S.R (1994) Evidence suggesting possible SCA1 gene involvement in schizophrenia. *Am J Hum Genet suppl.* SS 867.

deLeon, J., Dadvand, M., Canuso, C., White, A.O., Stanilla, J.K., and Simpson, G.M. (1995): Schizophrenia and smoking: an epidemiological survey in a state hospital. *Am.J.Psychiatry*, 152(3):453-455.

DeLisi, L.E. and Lieberman, J. (1991): Longitudinal perspectives on the pathophysiology of schizophrenia - examining the neurodevelopmental versus neurodegenerative hypotheses. *Schizophrenia Research*, 5:183-210.

DeLisi, L.E. (1992): The significance of age of onset for schizophrenia. *Schizophr.Bull.*, 18:209-215.

DeLisi, L.E., Friedrich, U., Wahlstrom, J., Boccio Smith, A., Forsman, A., Eklund, K., and Crow, T.J. (1994): Schizophrenia and sex chromosome anomalies.

Schizophr.Bull., 20:495-505.

DeLisi, L.E., Devoto, M., Lofthouse, R., Poulter, M., Smith, A., Shields, G., Bass, N., Chen, G., Vita, A., Morganti, C., Ott, J., and Crow, T.J. (1994): Search for linkage to schizophrenia on the X and Y chromosomes. *Am.J.Med.Genet.*, 54:113-121.

Derisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y.D., Su, Y.A., and Trent, J.M. (1996): Use of a cdna microarray to analyze gene-expression patterns in human cancer. *Nature Genetics*, 14:457-460.

Devereux, J., Haeberli, P., and Smithies, O. (1984): A comprehensive set of sequence analysis programs for the VAX. *Nucl.Acids Res.*, 12:387-395.

Devlin, B., Daniels, M., and Roeder, K. (1997): The heritability of IQ. *Nature*, 388:468-471.

Devon, R.S., Millar, J.K., Anderson, S., Christie, S., Maule, J.C., Shibasaki, Y., Evans, K.L., Brown, J., WilsonAnnan, J.C., Lawson, D., Gosden, J.R., Muir, W.J., Blackwood D.H.R., Stclair, D.M., Brookes, A.J., and Porteous, D.J. (1997): Molecular genetic analysis of a translocation breakpoint associated with schizophrenia. *Schizophrenia Research*, 24:56-56.

Devon, R.S., Evans, K.L., Maule, J.C., Christie, S., Anderson, S., Brown, J., Shibasaki, Y., Porteous, D.J., and Brookes, A.J. (1997): Novel transcribed sequences neighbouring a translocation breakpoint associated with schizophrenia. *American Journal Of Medical Genetics*, 74:82-90.

Dietrich, W.F., Lander, E.S., Smith, J.S., Moser, A.R., Gould, K.A., Luongo, C., Borenstein, N., and Dove, W. (1993): Genetic identification of *Mom-1*, a major modifier locus affecting min-induced intestinal neoplasia in the mouse. *Cell*, 75:631-639.

Dohan, F.C. (1982): Response to gluten versus neuronal damage. *American Journal of Psychiatry*, 139:1376-1376.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., and Mattick, J.S. (1991): "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucl.Acids Res.*, 19:4008-4008.

Done, D.J., Crow, T.J., Johnstone, E.C., and Sacker, A. (1994): Childhood antecedents of schizophrenia and affective illness: social adjustment at ages 7 and 11. *BMJ.*, 309:699-703.

Durany, N., Thome, J., Palomo, A., Foley, P., Riederer, P., and CruzSanchez, F.F. (1996): Homozygosity at the dopamine D3 receptor gene in schizophrenic patients. *Neuroscience Letters*, 220:151-154.

Duyk, G.M., Kim, S., Myers, R.M., and Cox, D.R. (1990): Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc.Natl.Acad.Sci.USA*, 87:8995-8999.

Eagles, J.M. (1991): Is schizophrenia disappearing. *British Journal of Psychiatry*, 158:834-835.

Eaton, W.W. (1991): Update on the epidemiology of schizophrenia. *Epidemiologic Reviews*, 13:320-328.

Ebstein, R.P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E.R., Nemanov, L., Katz, M., and Belmaker, R.H. (1996): Dopamine d4 receptor (D4dr) Exon-iii polymorphism associated with the human personality-trait of novelty seeking. *Nature Genetics*, 12:78-80.

Ebstein, R.P., Osher, Y., and Belmaker, R.H. (1996): D4dr exon-iii polymorphism association with personality variation in normals. *Biological Psychiatry*, 39:222-

222.

Endicott, J. and Spitzer, R.L. (1978): A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry*, 35:837-844.

Erdmann, J., Shimron-Abarbanell, D., Cichon, S., Albus, M., Maier, W., Lichtermann, D., Minges, J., Reuner, U., Franzek, E., Ertl, M.A., Hebebrand, J., Remschmidt, H., Lehmkuhl, G., Poustka, F., Schmidt, M., Fimmers, R., Korner, J., Rietschel, M., Propping, P., and Nothen, M.M. (1995): Systematic screening for mutations in the promoter and the coding region of the 5-HT_{1A} gene. *Am.J.Med.Gen.*, 60:393-399.

Erlenmeyer-Kimling, L. and Paradowski, W. (1966): Selection and schizophrenia. *The American Naturalist*, 100:651-665.

Evans, K.L., Brown, J., Shibasaki, Y., Devon, R.S., He, L., Arveiler, B., Christie, S., Maule, J.C., Baillie, D., Slorach, E.M., Anderson, S.M., Gosden, J.R., Petit, J., Weith, A., Gosden, C.M., Blackwood, D.H.R., St.Clair, D.M., Muir, W.J., Brookes, A.J., and Porteous, D.J. (1995): A contiguous clone map over 3Mb on the long arm of chromosome 11 across a balanced translocation associated with schizophrenia. *Genomics*, 28:420-428.

Evans, K. L. Mapping a balanced translocation t(1;11)(q42.2;q21) linked to schizophrenia. 1997. University of Edinburgh. 1993.

Ewens, W.J. and Spielman, R.S. (1995): The transmission disequilibrium test - history, subdivision and admixture. *American Journal Of Human Genetics*, 57:455-464.

Fahy, T.A., Jones, P.B., Sham, P.C., and Murray, R.M. (1992): Schizophrenia in afro-caribbeans in the uk following prenatal exposure to the 1957 a2 influenza epidemic. *Schizophrenia Research*, 6:98-99.

Falk, C.T. and Rubinstein, P. (1987): Haplotype relative risks - an easy reliable way to construct a proper control sample for risk calculations. *Annals Of Human Genetics*, 51:227-233.

Falkai, P., Bogerts, B., and Rozumek, M. (1988): Limbic pathology in schizophrenia: the enteorhinal region-a morphometric study. *Biol.Psychiatry.*, 24:515-521.

Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., Mannens, M., Danes, S., vanHeyningen, V., and Hanson, I. (1995): Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Human Molecular Genetics*, 4:415-422.

Faraone, S.V. and Tsuang, M.T. (1985): Quantitative models of the genetic transmission of schizophrenia. *Psychological Bulletin*, 98(1):41-66.

Faraone, S.V., Chen, W.J., Goldstein, J.M., and Tsuang, M.T. (1994): Gender differences in age at onset of schizophrenia. *British Journal of Psychiatry*, 164:625-629.

Farmer, A.E., Williams, J., and Jones, I. (1994): Phenotypic definitions of psychotic illness for molecular genetic research. *Am.J.Med.Genet.*, 54:365-371.

Feinberg, A.P. and Vogelstein, A. (1984): A technique for radiolabeling DNA fragments to high specific activity. *Anal.Biochem.*, 137:266-267.

Fischer, M., Harvald, B., and Hauge, M. (1969): A Danish twin study of schizophrenia. *Br.J.Psychiatry*, 115:981-990.

Fischer, M. (1971): Psychosis in the offspring of schizophrenia monozygotic twins and their normal co-twins. *Br.J.Psychiatry*, 118:43-52.

Fletcher, J.M., Evans, K., Baillie, D., Byrd, P., Hanratty, D., Leach, S., Julier, C., Gosden, J.R., Muir, W., Porteous, D.J., St.Clair, D., and van Heyningen, V. (1993): Schizophrenia-associated chromosome 11q21 translocation - identification of flanking markers and development of chromosome 11q fragment hybrids as cloning and mapping resources. *American Journal Of Human Genetics*, 52:478-490.

Flint, J., Corley, R., Defries, J.C., Fulker, D.W., Gray, J.A., Miller, S., and Collins, A.C. (1995): A simple genetic-basis for a complex psychological trait in laboratory mice. *Science*, 269:1432-1435.

Fogelson, D.L., Nuechterlein, K.H., Asarnow, R.F., Subotnik, K.L., and Talovic, S.A. (1991): Interrater reliability of the Structured Clinical Interview for DSM-III- R, Axis II: schizophrenia spectrum and affective spectrum disorders. *Psychiatry.Res.*, 39:55-63.

Forscher, P. and Smith, S.J. (1988): Actions of cytochalasins on the organization of actin-filaments and microtubules in a neuronal growth cone. *Journal of Cell Biology*, 107:1505-1516.

Freedman, R., Coon, H., MylesWorsley, M., OrrUrtreger, A., Olincy, A., Davis, A., Polymeropoulos, M., Holik, J., Hopkins, J., Hoff, M., Rosenthal, J., Waldo, M.C., Reimherr, F., Wender, P., Yaw, J., Young, D.A., Breese, C.R., Adams, C., Patterson, D., Adler, L.E., Kruglyak, L., Leonard, S., and Byerley, W. (1997): Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 94:587-592.

Freedman, R.F., Adler, L., Bickford, P., Byerley, W., Coon, H., Cullum, C.M., Griffin, J.M., Harris, J.G., Leonard, S., Miller, C., Myles-Worsley, M., Nagamoto, H.T., Rose, G., and Waido, M. (1994): Schizophrenia and nicotinic receptors. *Harvard Review of Psychiatry*, 2:179-192.

Gallant, D.M. (1990): Diagnosis of the schizophrenic disorders. *Psychiatr.Med.*, 8:21-40.

Gattaz, W.F., Hubner, C.V., Nevalainen, T.J., Thuren, T., and Kinnunen, P.K. (1990): Increased serum phospholipase A2 activity in schizophrenia: a replication study. *Biol.Psychiatry*, 28:495-501.

Gershon, E.S. and Rieder, R.O. (1992): Major disorders of mind and brain. *Sci.Am.*, 267:126-133.

Gibson, F., Walsh, J., Mburu, P., Varela, A., Brown, K.A., Antonio, M., Beisel, K.W., Steel, K.P., and Brown, S.D.M. (1995): A type-vii myosin encoded by the mouse deafness gene shaker-1. *Nature*, 374:62-64.

Gill, M., Vallada, H., Collier, D., Sham, P., Holmans, P., Murray, R., McGuffin, P., Nanko, S., Owen, M., Antonarakis, S., Housman, D., Kazazian, H., Nestadt, G., Pulver, A.E., Straub, R.E., Maclean, C.J., Walsh, D., Kendler, K.S., DeLisi, L., Polymeropoulos, M., Coon, H., Byerley, W., Lofthouse, R., Gershon, E., Golden, L., Crow, T., Freedman, R., Laurent, C., Bodeaupean, S., Damato, T., Jay, M., Campion, D., Mallet, J., Wildenauer, D.B., Lerer, B., Albus, M., Ackenheil, M., Ebstein, R.P., Hallmayer, J., Maier, W., Gurling, H., Curtis, D., Kalsi, G., Brynjolfsson, J., Sigmundson, T., Petursson, H., Blackwood, D., Muir, W., Stclair, D., He, L., Maguire, S., Moises, H.W., Hwu, H.G., Yang, L., Wiese, C., Tao, L., Liu, X.H., Kristbjarnason, H., Levinson, D.F., Mowry, B.J., Doniskeller, H., Hayward, N.K., Crowe, R.R., Silverman, J.M., Nancarrow, D.J., and Read, C.M. (1996): A combined analysis of d22s278 marker alleles in affected sib-pairs - support for a susceptibility locus for schizophrenia at chromosome 22q12. *American Journal Of Medical Genetics*, 67:40-45.

Gilmore, J.H., Perkins, D.O., Kliewer, M.A., Hage, M.L., Silva, S.G., Chescheir, N.C., Hertzberg, B.S., and Sears, C.A. (1996): Fetal brain-development of twins assessed in-utero by ultrasound - implications for schizophrenia. *Schizophrenia*

Research, 19:141-149.

Gorwood, P., Leboyer, M., Damato, T., Jay, M., Campion, D., Hillaire, D., Mallet, J., and Feingold, J. (1992): Evidence for a pseudoautosomal locus for schizophrenia .1. a replication study using phenotype analysis. *British Journal of Psychiatry*, 161:55-58.

Gorwood, P., Leboyer, M., Jay, M., Payan, C., and Feingold, J. (1995): Gender and age at onset in schizophrenia - impact of family history. *American Journal of Psychiatry*, 152:208-212.

Gorwood, P., Leboyer, M., Falissard, B., Jay, M., Rouillon, F., and Feingold, J. (1996): Anticipation in schizophrenia - new light on a controversial problem. *American Journal of Psychiatry*, 153:1173-1177.

Gottesman, I.I. and Wolfgram, D.L. (1991): *Schizophrenia Genesis: The Origins of Madness*, AnonymousFreeman, W.H., New York.

Gottesman, I.I. (1991): Genetic epidemiology and models for schizophrenia. *American Journal Of Human Genetics*, 49:26-26.

Gottesman, I.I. (1997): Human genetics - Twins: En route to QTLs for cognition. *Science*, 276:1522-1523.

Green, E.D., Mohr, R.M., Idol, J.R., Jones, M., Buckingham, J.M., Deaven, L.L., Moyzis, R.K., and Olson, M.V. (1991): Systematic generation of sequence-tagged sites for physical mapping of human-chromosomes - application to the mapping of human chromosome-7 using yeast artificial chromosomes. *Genomics*, 11:548-564.

Green, E.D. and Waterston, R.H. (1991): The human genome project - prospects and implications for clinical medicine. *Jama-Journal Of The American Medical*

Association, 266:1966-1975.

Green, M.F., Satz, P., Smith, C., and Nelson, L. (1989): Is there atypical handedness in schizophrenia? *J.Abnorm.Psychol.*, 98:57-61.

Gurling, H., Kalsi, G., Chen, A.H.S., Green, M., Butler, R., Read, T., Murphy, P., Curtis, D., Sharma, T., Petursson, H., and Brynjolfsson, T. (1995): Schizophrenia susceptibility and chromosome 6p24-22. *Nature Genetics*, 11:234-235.

Hacia, J.G., Brody, L.C., Chee, M.S., Fodor, S.A., and Collins, F.S. (1996): Detection of heterozygous mutations in BRCA1 using high-density oligonucleotide arrays and 2-color fluorescence analysis. *Nature Genetics*, 14:441-447.

Harding, C.M., Brooks, G.W., Ashikaga, T., Strauss, J.S., and Breier, A. (1987): The vermont longitudinal-study of persons with severe mental-illness .2. Long-term outcome of subjects who retrospectively met dsm-iii criteria for schizophrenia. *American Journal of Psychiatry*, 144:727-735.

Hare, E.H. (1987): Seasonal incidence and secular trends in schizophrenia. *International Journal Of Neuroscience*, 32:477-477.

Harrison, G., Owens, D., Holton, D., Nielson, D., and Boot, D. (1988): Schizophrenia in the UK Afro-Caribbean population. *Schizophrenia Research*, 1:119-119.

Harrison, G. and Mason, P. (1993): Schizophrenia--falling incidence and better outcome? *Br.J.Psychiatry*, 163:535-541.

Haverkamp, F., Propping, P., and Hilger, T. (1982): Is there an increase of reproductive rates in schizophrenia? *Arch Psych Neurol Sci*, 232:439-450.

He, L., Mansfield, D.C., Brown, A.F., Green, D.K., Morris, S.W., Stclair, D.M., Muir, W.J., Maclean, A., Wright, A.F., and Blackwood, D.R. (1995): Automated linkage

analysis in psychiatric-disorders. *American Journal Of Medical Genetics*, 60:192-198.

He, L., Mansfield, D.C., Brown, A.F., Green, D.K., Morris, S.W., St.Clair, D.M., Muir, W.J., Maclean, A., Wright, A.F., and Blackwood, D.H.R. (1995): Automated linkage analysis in psychiatric disorders. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 60:192-198.

Hegarty, J.D., Baldessarini, R.J., Tohen, M., Waternaux, C., and Oepen, G. (1994): One hundred years of schizophrenia: a meta-analysis of the outcome literature. *Am.J.Psychiatry*, 151:1409-1416.

Hemmings, G. (1990): Causes of schizophrenia. *Nutr.Health*, 7:11-19.

Heston, L.L. (1966): Psychiatric disorders in foster home reared children of schizophrenic mothers. *Br.J.Psychiatry*, 112:819-825.

Heyman, I. and Murray, R.M. (1992): Schizophrenia and neurodevelopment. *Journal Of The Royal College Of Physicians Of London*, 26:143-146.

Hilbert, P., Lindpaintner, K., Beckmann, J.S., Serikawa, T., Soubrier, F., Dubay, C., Cartwright, P., Degouyon, B., Julier, C., Takahasi, S., Vincent, M., Ganten, D., Georges, M., and Lathrop, G.M. (1991): Chromosomal mapping of 2 genetic-loci associated with blood-pressure regulation in hereditary hypertensive rats. *Nature*, 353:521-529.

Holden, R.J., Mooney, P.A., and Newman, J.C. (1994): Schizophrenia: an extended etiological explanation. *Med.Hypotheses.*, 42:115-123.

Holland, G. and Gosden, C. (1990): A balanced chromosomal translocation partially segregating with psychotic illness in a family. *Psych.Res.*, 32:1-8.

Holland, T. and Gosden, C. (1990): A balanced chromosomal translocation partially co-segregating with psychotic illness in a family. *Psychiatry Research*, 32:1-8.

Holzman, P.S. (1992): Behavioral markers of schizophrenia useful for genetic studies. *J.Psychiatr.Res.*, 26:427-445.

Horrobin, D.F. and Huang, Y.-S. (1983): Schizophrenia: the role of abnormal essential fatty acid and prostaglandin metabolism. *Medical Hypotheses*, 10:329-336.

Horrobin, D.F., Glen, A.I., and Vaddadi, K. (1994): The membrane hypothesis of schizophrenia. *Schizophr.Res.*, 13:195-207.

Huntington's Disease Collaborative Research Group (1993): A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, 72:971-983.

International classification of disease 10 (ICD10) Classification of mental and behavioural disorders Geneva: World Health Organisation 1992

Inayama, Y., Yoneda, H., Sakai, T., Ishida, T., Nonomura, Y., Kono, Y., Takahata, R., Koh, J., Sakai, J., Takai, A., Inada, Y., and Asaba, H. (1996): Positive association between a dna-sequence variant in the serotonin 2a receptor gene and schizophrenia. *American Journal Of Medical Genetics*, 67:103-105.

Ingvar, D.H. and Franzen, G. (1997): Abnormalities of cerebral blood flow distribution in patients with chronic schizophrenia. *Acta Psychiatr.Scand.*, 50:462

Jakob, H. and Beckmann, H. (1986): Prenatal developmental disturbances in the limbic allocortex in schizophrenics. *J.Neural.Transm.*, 65:303-326.

Jentsch, J.D., Redmond, D.E., Elsworth, J.D., Taylor, J.R., Youngren, K.D., and Roth, R.H. (1997): Enduring cognitive deficits and cortical dopamine dysfunction

in monkeys after long-term administration of phencyclidine. *Science*, 277:953-955.

Johnstone, E.C., Crow, T.J., Frith, D.C., Husband, J., and Krel, L. (1976): Cerebral ventricular size and cognitive impairment in schizophrenia. *Lancet*, ii:924-924.

Johnstone, E.C. (1993): Schizophrenia: problems in clinical practice. *Lancet*, 341:536-538.

Jones, C.T., Morris, S., Yates, C.M., Moffoot, A., Sharpe, C., Brock, D.J.H., and St Clair, D. (1992): Mutation in codon 713 of the B amyloid precursor protein gene presenting with schizophrenia. *Nature Genetics*, 1:306-309.

Jones, D.S.C. and Schofield, J.P. (1990): A rapid method for isolating high quality plasmid DNA suitable for DNA sequencing. *N.A.R.*, 18:7463-7464.

Jones, D.S.C. and Schofield, J.P. (1990): A rapid method for isolating high quality plasmid DNA suitable for DNA sequencing. *Nucl.Acids Res.*, 18:7463-7464.

Jones, P., Rodgers, B., Murray, R., and Marmot, M. (1994): Child development risk factors for adult schizophrenia in the British 1946 birth cohort. *Lancet*, 344:1398-1402.

Jonsson, E., Lannfelt, L., Sokoloff, P., Schwartz, J.C., and Sedvall, G. (1993): Lack of association between schizophrenia and alleles in the dopamine-d3 receptor gene. *Acta Psychiatrica Scandinavica*, 87:345-349.

Jonsson, E., Nothen, MM., Bunzel, R., Propping, P., Sedvall, G. (1996):Letter *Lancet* 347:1831

Joos, S., Haluska, F.G., Falk, M.H., Henglein, B., Hameister, H., Croce, C.M., and Bornkamm, G.W. (1992): Mapping chromosomal breakpoints of burkitts t(8-14)

Translocations far upstream of c-myc. *Cancer Research*, 52:6547-6552.

Julier, C., Hyer, R.N., Davies, J., Merlin, F., Soularue, P., Briant, L., Cathelineau, G., Deschamps, I., Rotter, J.I., Froguel, P., Boitard, C., Bell, J.I., and Lathrop, G.M. (1991): Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. *Nature*, 354:155-159.

Jurka, J., Walichiewicz, J., and Milosavljevic, A. (1992): Prototypic sequences for human repetitive dna. *Journal Of Molecular Evolution*, 35:286-291.

Kalsi, G., Brynjolfsson, J., Butler, R., Sherrington, R., XX, Sigmundsson, T., Read, T., Murphy, P., Sharma, T., Petursson, H., and Gurling, H.M.D. (1995): Linkage analysis of chromosome 22q12-13 in a United Kingdom/Icelandic sample of 23 multiplex schizophrenia families. *Am.J.Med.Gen.*, 60:298-301.

Keefe, R.S., Silverman, J.M., Siever, L.J., and Cornblatt, B.A. (1991): Refining phenotype characterization in genetic linkage studies of schizophrenia. *Soc.Biol.*, 38:197-218.

Kendell, R.E. (1987): Diagnosis and classification of functional psychoses. *British Medical Bulletin*, 43:499-513.

Kendell, R.E. and Kemp, I.W. (1989): Maternal influenza in the etiology of schizophrenia. *Arch.Gen.Psychiatry*, 46:878-882.

Kendell, R.E., Malcolm, D.E., and Adams, W. (1993): The problem of detecting changes in the incidence of schizophrenia. *British Journal of Psychiatry*, 162:212-218.

Kendler, K.S. and Gruenberg, A.M. (1984): An independent analysis of the danish adoption study of schizophrenia .6. The relationship between psychiatric-disorders as defined by dsm- iii in the relatives and adoptees. *Archives Of*

General Psychiatry, 41:555-564.

Kendler, K.S., Gruenberg, A.M., and Tsuang, M.T. (1985): Psychiatric illness in first degree relatives of schizophrenic and surgical control patients. *Arch.Gen.Psychiatry*, 42:770-779.

Kennedy, J.L., Giuffra, L.A., Moises, H.W., Cavalli-Sforza, L.L., Pakstis, A.J., Kidd, J.R., Casriglione, C.M., Sjogren, B., Wetterburg, L., and Kidd, K.K. (1988): Evidence against linkage of schizophrenia to markers on chromosome 5 in a Northern Swedish pedigree. *Nature*, 336:167-170.

Kerckaert, J.P., Deweindt, C., Tilly, H., Quief, S., Lecocq, G., and Bastard, C. (1993): Laz3, a novel zinc-finger encoding gene, is disrupted by recurring chromosome-3q27 translocations in human lymphomas. *Nature Genetics*, 5:66-70.

Kerwin, R.W. (1993): Glutamate receptors, microtubule associated proteins and developmental anomaly in schizophrenia: an hypothesis [editorial]. *Psychol Med*, 23:547-551.

Kety, S.S., Rosenthal, D., Wender, P.H., and Schulsinger, F. (1971): Mental illness in the biological and adoptive relatives of adopted schizophrenics. *Am.J.Psychiatry*, 128:301-306.

Kety, S.S., Rosenthal, D., Wender, P.H., Schulsinger, F., and Jacobson, B. (1978): The biologic and adoptive families of individuals who become schizophrenic: prevalance of mental illness and other characteristics. In: *The Nature of Schizophrenia*, edited by L.C. Wynne, et al, pp. 25-37. John Wiley and Sons Inc., New York.

Kety, S.S. (1988): Schizophrenic illness in the families of schizophrenic adoptees - findings from the danish national sample. *Schizophrenia Bulletin*, 14:217-222.

King, M., Coker, E., Leavey, G., Hoare, A., and Johnson Sabine, E. (1994): Incidence of psychotic illness in London: comparison of ethnic groups. *BMJ.*, 309:1115-1119.

Knight, J., Knight, A., and Ungvari, G. (1992): Can autoimmune mechanisms account for the genetic predisposition to schizophrenia? *Br.J.Psychiatry.*, 160:533-540.

Kohn, M.L. (1976): The interaction of social class and other factors in the etiology of schizophrenia. *Am.J.Psychiat.*, 133:177-180.

Kotler, M., Cohen, H., Segman, R., Gritsenko, I., Nemanov, L., Lerer, B., Kramer, I., ZerZion, M., Kletz, I., and Ebstein, R.P. (1997): Excess dopamine D4 receptor (D4DR) exon III seven repeat allele in opioid-dependent subjects. *Molecular Psychiatry*, 2:251-254.

Kraeplin, E Dementia praecox and paraphrenia Barclay RM trans. Edinburgh: E+S Livingstone 1919

Kringlen, E. (1967): *Heredity and environment in the functional psychoses.* Heinemann Medical Books, London.

Kringlen, E. (1993): Genes and environment in mental-illness - perspectives and ideas for future-research. *Acta Psychiatrica Scandinavica*, 87:79-84.

Krizman, D.B. and Berget, S.M. (1993): Efficient selection of 3'-terminal exons from vertebrate dna. *Nucleic Acids Research*, 21:5198-5202.

Labuda, M.C., Gottesman, I.I., and Pauls, D.L. (1993): Usefulness of twin studies for exploring the etiology of childhood and adolescent psychiatric-disorders. *American Journal Of Medical Genetics*, 48:47-59.

Lander, E. and Kruglyak, L. (1995): Genetic dissection of complex traits - guidelines for interpreting and reporting linkage results. *Nature Genetics*, 11:241-247.

Lander, E.S. and Schork, N.J. (1994): Genetic dissection of complex traits. *Science*, 265:2037-2048.

Lapensee, M.A. (1992): A review of schizoaffective disorder: I. Current concepts. *Canadian Journal of Psychiatry*, 37:335-346.

Laurent, C., Savoye, C., Samolyk, D., Meloni, R., Mallet, J., Campion, D., Martinez, M., Damato, T., Bastard, C., and Dollfus, S. (1994): Homozygosity at the dopamine d3 receptor locus is not associated with schizophrenia. *Journal Of Medical Genetics*, 31:260-260.

Leboyer, M., Filteau, M.J., Jay, M., Campion, D., Rochet, T., Damato, T., Feingold, J., Deslauriers, A., and Widlocher, D. (1992): No gender effect on age at onset in familial schizophrenia. *American Journal of Psychiatry*, 149:1409-1409.

Leff, J., Sartorius, N., Jablensky, A., Korten, A., and Ernberg, G. (1992): The International Pilot Study of Schizophrenia: five-year follow- up findings. *Psychol.Med.*, 22:131-145.

Leff, J. (1992): Over the edge: stress and schizophrenia. *New Scientist, 4th Jan*, 30-33.

Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M.B. (1996): The image consortium - an integrated molecular analysis of genomes and their expression. *Genomics*, 33:151-152.

Letourneau, P.C. and Shattuck, T.A. (1989): Distribution and possible interactions of actin-associated proteins and cell-adhesion molecules of nerve growth cones. *Development*, 105:505-519.

Levinson, D.F., Wildenauer, D.B., Schwab, S.G., Albus, M., Hallmayer, J., Lerer, B., Maier, W., Blackwood, D., Muir, W., Stclair, D., Morris, S., Moises, H.W., Yang, L., Kristbjarnarson, H., Helgason, T., Wiese, C., Collier, D.A., Holmans, P., Daniels, J., Rees, M., Asherson, P., Roberts, Q., Cardno, A., Arranz, M.J., Vallada, H., Mcguffin, D., Owen, M.J., Pulver, A.E., Antonarakis, S.E., Babb, R., Blouin, J.L., Demarchi, N., Dombroski, B., Housman, D., Karayiorgou, M., Ott, J., Kasch, L., Kazazian, H., Lasseter, V.K., Loetscher, E., Luebbert, H., Nestadt, G., Ton, C., Wolyniec, P.S., Laurent, C., Dechaldee, M., Thibaut, F., Jay, M., Samolyk, D., Petit, M., Campion, D., Mallet, J., Straub, R.E., Maclean, C.J., Easter, S.M., Oneill, F.A., Walsh, D., Kendler, K.S., Gejman, P.V., Cao, Q.H., Gershon, E., Badner, J., Beshah, E., Zhang, J., Riley, B.P., Rajagopalan, S., Mogudicarter, M., Jenkins, T., Williamson, R., DeLisi, L.E., Garner, C., Kelly, M., Leduc, C., Cardon, L., Lichter, J., Harris, T., Loftus, J., Shields, G., Comasi, M., Vita, A., Smith, A., Dann, J., Joslyn, G., Gurling, H., Kalsi, G., Brynjolfsson, J., Curtis, D., Sigmundsson, T., Butler, R., Read, T., Murphy, P., Chen, A.C.H., Petursson, H., Byerley, B., Hoff, M., Holik, J., Coon, H., Nancarrow, D.J., Crowe, R.R., Andreasen, N., Silverman, J.M., Mohs, R.C., Siever, L.J., Endicott, J., Sharpe, L., Walters, M.K., Lennon, D.P., Hayward, N.K., Sandkuijl, L.A., Mowry, B.J., Aschauer, H.N., Meszaros, K., Lenzinger, E., Fuchs, K., Heiden, A.M., Kruglyak, L., Daly, M.J., and Matise, T.C. (1996): Additional support for schizophrenia linkage on chromosome-6 and chromosome-8 - a multicenter study. *American Journal Of Medical Genetics*, 67:580-594.

Lewis, G., David, A., Andreasson, S., and Allebeck, P. (1992): Schizophrenia and city life. *Lancet*, 340:137-140.

Li, T., Yang, L., Wiese, C., Xu, C.T., Zeng, Z., Giros, B., Caron, M.G., Moises, H.W., and Liu, X.H. (1994): No association between alleles or genotypes at the dopamine transporter gene and schizophrenia. *Psychiatry Research*, 52:17-23.

Lijam, N., Paylor, R., McDonald, M.P., Crawley, J.N., Deng, C.X., Herrup, K., Stevens, K.E., Maccaferri, G., Mcbain, C.J., Sussman, D.J., and WynshawBoris, A. (1997): Social interaction and sensorimotor gating abnormalities in mice

lacking Dvl1. *Cell*, 90:895-905.

Lim, L.C., Nothen, M.M., Korner, J., Rietschel, M., Castle, D., Hunt, N., Propping, P., Murray, R., and Gill, M. (1994): No evidence of association between dopamine D4 receptor variants and bipolar affective disorder. *Am.J.Med.Genet.*, 54:259-263.

Lisitsyn, N. and Wigler, M. (1993): Cloning the differences between two complx genomes. *Science.*, 259:946-951.

Lisitsyn, N. and Wigler, M. (1995): Representational difference analysis in detection of genetic lesions in cancer. *Methods in Enzymology*, 254:291-304.

Lisitsyn, N.A., Leach, F.S., Vogelstein, B., and Wigler, M.H. (1994): Detection of genetic loss in tumors by representational difference analysis. *Cold Spring Harbor Symposia On Quantitative Biology*, 59:585-587.

Lisitsyn, N.A. (1995): Representational difference analysis - finding the differences between genomes. *Trends in Genetics*, 11:303-307.

Lohr, J.B. and Bracha, H.S. (1989): Can schizophrenia be related to prenatal exposure to alcohol? Some speculations. *Schizophr.Bull.*, 15(4):595-603.

Lovett, M., Kere, J., and Hinton, L.M. (1991): Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. *Proc.Natl.Acad.Sci.USA*, 88:9628-9632.

Lovett, M. (1994): Fishing for complements: finding genes by direct selection. *Trends Genet.*, 10:352-357.

Luongo, C., Gould, K.A., Su, L.K., Kinzler, K.W., Vogelstein, B., Dietrich, W., Lander, E.S., and Moser, A.R. (1993): Mapping of multiple intestinal neoplasia

(min) to proximal chromosome-18 of the mouse. *Genomics*, 15:3-8.

Macciardi, F., Kennedy, J.L., Ruocco, L., Giuffra, L., Carrera, P., Marino, C., Rinaldi, V., Smeraldi, E., and Ferrari, M. (1992): A genetic linkage study of schizophrenia to chromosome 5 markers in a Northern Italian population. *Biol.Psychiatry*, 31:720-728.

Macciardi, F., Verga, M., Kennedy, J.L., Petronis, A., Bersani, G., Pancheri, P., and Smeraldi, E. (1994): An association study between schizophrenia and the dopamine receptor genes DRD3 and DRD4 using haplotype relative risk. *Hum.Hered.*, 44:328-336.

Macphee, M., Chepenik, K.P., Liddell, R.A., Nelson, K.K., Siracusa, L.D., and Buchberg, A.M. (1995): The secretory phospholipase-a2 gene is a candidate for the mom1 locus, a major modifier of apc(min)-induced intestinal neoplasia. *Cell*, 81:957-966.

Maddox, J. (1997): Schizophrenia - The price of language? *Nature*, 388:424-425.

Maier, W., Minges, J., Eckstein, N., Brodski, C., Albus, M., Lerer, B., Hallmayer, J., Fimmers, R., Ackenheil, M., Ebstein, R.E., Borrmann, M., Lichtermann, D., and Wildenauer, D.B. (1996): Genetic-relationship between dopamine transporter gene and schizophrenia - linkage and association. *Schizophrenia Research*, 20:175-180.

Malhotra, A.K., Pinals, D.A., Weisenfeld, N., Pickar, D., and Breier, A. (1996): Nmda receptor function and schizophrenia - studies with ketamine. *Biological Psychiatry*, 39:296-296.

Malhotra, A.K., Goldman, D., Buchanan, R., Breier, A., and Pickar, D. (1996): 5Ht(2A) Receptor t102c polymorphism and schizophrenia. *Lancet*, 347:1830-1831.

Mansfield, D.C., Brown, A.F., Green, D.K., Carothers, A.D., Morris, S.W., Evans, H.J., and Wright, A.F. (1994): Automation of genetic-linkage analysis using fluorescent microsatellite markers. *Genomics*, 24:225-233.

Mant, R., Williams, J., Asherson, P., Parfitt, E., McGuffin, P., and Owen, M.J. (1994): Relationship between homozygosity at the dopamine D3 receptor gene and schizophrenia. *Am.J.Med.Genet.*, 54:21-26.

Matsudaira, P. (1991): Modular organization of actin cross-linking proteins. *Trends In Biochemical Sciences*, 16:87-92.

Mcgue, M. and Gottesman, I.I. (1989): Genetic-linkage in schizophrenia - perspectives from genetic epidemiology. *Schizophrenia Bulletin*, 15:453-464.

McGuffin, P., Reveley, A., and Holland, A. (1982): Identical triplets: non-identical psychosis? *Br.J.Psychiatry*, 140:1-6.

McGuffin, P., Owen, M.J., Asherson, P., and Farmer, A.E. (1994): Genetics, chance and dysmorphogenesis in schizophrenia. *British Journal of Psychiatry*, 165:694-695.

McGuffin, P., Asherson, P., Owen, M., and Farmer, A. (1994): The strength of the genetic effect - is there room for an environmental influence in the etiology of schizophrenia. *British Journal of Psychiatry*, 164:593-599.

Mednick, S.A., Machon, R.A., Huttunen, M.O., and Bonett, D. (1988): Adult schizophrenia following prenatal exposure to an influenza epidemic. *Arch.Gen.Psychiatry*, 45:189-192.

Meisler, M.H. (1996): The role of the laboratory mouse in the human genome project. *American Journal Of Human Genetics*, 59:764-771.

Moises, H.W., Gelernter, J., Giuffra, L.A., Zarcone, V., Wetterberg, L., Civelli, O., Kidd, K.K., Cavallisforza, L.L., Grandy, D.K., Kennedy, J.L., Vinogradov, S., Mauer, J., Litt, M., and Sjogren, B. (1991): No linkage between d2 dopamine receptor gene region and schizophrenia. *Archives Of General Psychiatry*, 48:643-647.

Moises, H.W., Yang, L., Kristbjarnarson, H., Wiese, C., Byerley, W., Macciardi, F., Arolt, V., Blackwood, D., Liu, X., and Sjogren, B. (1995): An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nat.Genet.*, 11:321-324.

Moises, H.W., Yang, L., Li, T., Havsteen, B., Fimmers, R., Baur, M.P., Liu, X.H., and Gottesman, I.I. (1995): Potential linkage disequilibrium between schizophrenia and locus d22s278 on the long arm of chromosome-22. *American Journal Of Medical Genetics*, 60:465-467.

Morris, A.G., Gaitonde, E., McKenna, P.J., Mollon, J.D., and Hunt, D.M. (1995): CAG repeat expansions and schizophrenia: association with disease in females and with early age-at-onset. *Hum.Mol.Genet.*, 4:1957-1961.

Moser, A.R., Dove, W.F., Roth, K.A., and Gordon, J.I. (1992): The min (multiple intestinal neoplasia) mutation - its effect on gut epithelial-cell differentiation and interaction with a modifier system. *Journal of Cell Biology*, 116:1517-1526.

Mount, S.M. (1982): A catalog of splice junction sequences. *Nucleic Acids Research*, 10:459-472.

Mowry, B.J., Nancarrow, D.J., Lennon, D.P., Sandkuijl, L.A., Crowe, R.R., Silverman, J.M., Mohs, R.C., Siever, L.J., Endicott, J., Sharpe, L., Walters, M.K., Hayward, N.K., and Levinson, D.F. (1995): Schizophrenia susceptibility and chromosome 6p24-22. *Nature Genetics*, 11:233-234.

MRC Field Review. Biological Psychiatry. 1993. MRC Field Review. (GENERIC) Ref Type: MRC Report

Muir, W.J., Gosden, C.M., Brookes, A.J., Fantes, J., Evans, K.L., Maguire, S.M., Stevenson, B., Boyle, S., Blackwood, D.H., St.Clair, D.M., Porteous, D.J., and Weith, A. (1995): Direct microdissection and microcloning of a translocation breakpoint region, t(1;11)(q42.2;q21), associated with schizophrenia. *Cytogenet.Cell Genet.*, 70:35-40.

Mullis, K.B. and Faloona, F.A. (1995): Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. *Methods Enzymol.*, 155:335-350.

Muramatsu, T., Higuchi, S., Murayama, M., Matsushita, S., and Hayashida, M. (1996): Association between alcoholism and the dopamine d4 receptor gene. *Journal Of Medical Genetics*, 33:113-115.

Murray, R.M., Ocallaghan, E., Castle, D.J., and Lewis, S.W. (1992): A neurodevelopmental approach to the classification of schizophrenia. *Schizophrenia Bulletin*, 18:319-332.

Murray, R.M., Jones, P., O'Callaghan, E., Takei, N., and Sham, P. (1992): Genes, viruses and neurodevelopmental schizophrenia. *J.Psychiatr.Res.*, 26:225-235.

Nanko, S., Gill, S., Owen, M., Takazawa, N., Moridaiwa, J., and Kazamatsuri, H. (1992): Linkage study of schizophrenia with markers on chromosome 11 in two Japanese pedigrees. *Jap.J.Psychiatr.Neurol.*, 46(1):155-159.

Nanko, S., Sasaki, T., Fukuda, R., Hattori, M., Dai, X.Y., Kazamatsuri, H., Kuwata, S., Juji, T., and Gill, M. (1993): A study of the association between schizophrenia and the dopamine D3 receptor gene. *Hum.Genet.*, 92:336-338.

Nanko, S., Fukuda, R., Hattori, M., Sasaki, T., Dai, X.Y., Yamaguchi, K., and Kazamatsuri, H. (1994): Further evidence of no linkage between schizophrenia

and the dopamine D3 receptor gene locus. *Am.J.Med.Genet.*, 54:264-267.

Neiswanger, K., Hill, S.Y., and Kaplan, B.B. (1995): Association and linkage studies of the taqi a1 allele at the dopamine d-2 receptor gene in samples of female and male alcoholics. *American Journal Of Medical Genetics*, 60:267-271.

Nelson, S.F., McCusker, J.H., Sander, M.A., Kee, Y., Modrich, P., and Brown, P.O. (1993): Genomic mismatch scanning : a new approach to genetic linkage mapping. *Nature Genet.*, 4,11-17.

Neylan, T.C. and van Kammen, D.P. (1990): Biological mechanisms of schizophrenia: an update. *Psychiatr.Med.*, 8:41-52.

NIH/CEPH Collaborative Mapping Group (1992): A comprehensive linkage map of the human genome. *Science*, 258:67-87.

Nimgaonkar, V.L., Zhang, X.R., Caldwell, J.G., Ganguli, R., and Chakravarti, A. (1993): Association study of schizophrenia with dopamine D3 receptor gene polymorphisms: probable effects of family history of schizophrenia? *Am.J.Med.Genet.*, 48:214-217.

Nimgaonkar, V.L., Sanders, A.R., Ganguli, R., Zhang, X.R., Brar, J., Hogge, W., Fann, W.E., Patel, P.I., and Chakravarti, A. (1996): Association study of schizophrenia and the dopamine d3 receptor gene locus in 2 independent samples. *American Journal Of Medical Genetics*, 67:505-514.

Nimgaonkar, V.L. (1997): In defense of genetic association studies. *Molecular Psychiatry*, 2:275-277.

North, K.N. and Beggs, A.H. (1996): Deficiency of a skeletal-muscle isoform of alpha-actinin (Alpha- actinin-3) In merosin-positive congenital muscular-dystrophy. *Neuromuscular Disorders*, 6:229-235.

Nothen, M.M., Cichon, S., Propping, P., Fimmers, R., Schwab, S.G., and Wildenauer, D.B. (1993): Excess of homozygosity at the dopamine D3 receptor gene in schizophrenia not confirmed [letter; comment]. *J.Med.Genet.*, 30:708-709.

Nothen, M.M., Cichon, S., Hemmer, S., Hebebrand, J., Remschmidt, H., Lehmkuhl, G., Poustka, F., Schmidt, M., Catalano, M., Fimmers, R., Korner, J., Rietschel, M., and Propping, P. (1994): Human dopamine D4 receptor gene: frequent occurrence of a null allele and observation of homozygosity. *Hum.Mol.Genet.*, 3:2207-2212.

O'Callaghan, E., Sham, P., Takei, N., Glover, G., and Murray, R.M. (1991): Schizophrenia after prenatal exposure to 1957 A2 influenza epidemic. *Lancet*, 337:1248-1249.

O'Callaghan, E., Gibson, T., Colohan, H.A., Buckley, P., Walshe, D.G., Larkin, C., and Waddington, J.L. (1992): Risk of schizophrenia in adults born after obstetric complications and their association with early onset of illness - a controlled-study. *British Medical Journal*, 305:1256-1259.

O'Callaghan, E., Sham, P.C., Takei, N., Murray, G., Glover, G., Hare, E.H., and Murray, R.M. (1994): The relationship of schizophrenic births to 16 infectious diseases. *Br.J.Psychiatry*, 165:353-356.

O'Donovan, M.C., Guy, C., Craddock, N., Murphy, K.C., Cardno, A.G., Jones, L.A., Owen, M.J., and McGuffin, P. (1995): Expanded CAG repeats in schizophrenia and bipolar disorder. *Nature Genetics*, 10:380-381.

O'Donovan, M.C., Guy, C., Craddock, N., Bowen, T., McKeon, P., Macedo, A., Maier, W., Wildenauer, D., Aschauer, H.N., Sorbi, S., Feldman, E., Mynettjohnson, L., Claffey, E., Nacmias, B., Valente, J., Dourado, A., Grassi, E., Lenzinger, E., Heiden, A.M., Moorhead, S., Harrison, D., Williams, J., McGuffin, P., and Owen, M.J. (1996): Confirmation of association between expanded

cag/ctg repeats and both schizophrenia and bipolar disorder. *Psychological Medicine*, 26:1145-1153.

O'Donovan, M.C. and Owen, M.J. (1996): Dynamic mutations and psychiatric genetics. *Psychological Medicine*, 26:1-6.

Okubo, Y., Suhara, T., Suzuki, K., Kobayashi, K., Inoue, O., Terasaki, O., Someya, Y., Sassa, T., Sudo, Y., Matsushima, E., Iyo, M., Tateno, Y., and Toru, M. (1997): Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET. *Nature*, 385:634-636.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989): Rapid and sensitive detection of point mutations and dna polymorphisms using the polymerase chain-reaction. *Genomics*, 5:874-879.

O'Rourke, D.H., Gottesman, I.I., Suarez, B.K., Rice, J., and Reich, T. (1982): Refutation of the general single-locus model for the etiology of schizophrenia. *American Journal Of Human Genetics*, 34:630-649.

Owen, F. and Simpson, M. (1994): The neurochemistry of schizophrenia. In: *Molecular and Cell Biology of Neuropsychiatric Diseases*, edited by F. Owen, et al, pp. 133-159. Chapman & Hall, London.

Owen, M.J., Mant, R., Parfitt, E., Williams, J., Asherton, P., Mahoney, G.O., Van Os, J., Llewellyn, D., Collier, D., and Gill, M. (1992): No Association between RFLPs at the porphobilinogen deaminase gene and schizophrenia. *Human Genetics*, 90:131-132.

Pallast, E.G.M., Jongbloet, P.H., Straatman, H.M., and Zielhuis, G.A. (1994): Excess seasonality of births among patients with schizophrenia and seasonal ovopathy. *Schizophrenia Bulletin*, 20:269-276.

Parimoo, S., Patanjali, S.R., Shukla, H., Chaplin, D.D., and Weissman, S.M. (1991): cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. *Proc.Natl.Acad.Sci.USA*, 88:9623-9627.

Paterson, A. (1997): Case-control association studies in complex traits - The end of an era? *Molecular Psychiatry*, 2:277-278.

Pearson, W.R. and Lipman, D.J. (1988): Improved tools for biological sequence comparison. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 85:2444-2448.

Penrose, L.S. (1953): The general purpose sib-pair linkage test. *Ann Eugen*, 18:120-124.

Perutz, M.F. (1996): Glutamine repeats and inherited neurodegenerative diseases: molecular aspects. *Curr.Opin.Struct.Biol.*, 6:848-858.

Peterson, A.C., Dirienzo, A., Lehesjoki, A.E., Delachapelle, A., Slatkin, M., and Freimer, N.B. (1995): The distribution of linkage disequilibrium over anonymous genome regions. *Human Molecular Genetics*, 4:887-894.

Petronis, A. and Kennedy, J.L. (1995): Unstable genes - unstable mind? *Am.J.Psychiatry*, 164-172.

Petronis, A., Bassett, A.S., Honer, W.G., Vincent, J.B., Tatuch, Y., Sasaki, T., Ying, D.J., Klempan, T.A., and Kennedy, J.L. (1996): Search for unstable dna in schizophrenia families with evidence for genetic anticipation. *American Journal Of Human Genetics*, 59:905-911.

Polymeropoulos, M.H., Coon, H., Byerley, W., Gershon, E.S., Goldin, L., Crow, T.J., Rubenstein, J., Hoff, M., Holik, J., Smith, A.M., Shields, G., Bass, N.J., Poulter, M., Lofthouse, R., Vita, A., Morganti, C., Merril, C.R., and DeLisi, L.E. (1994): Search for a schizophrenia susceptibility locus on human chromosome 22.

Am.J.Med.Genet., 54:93-99.

Pope, H.G., Jonas, J.M., Cohen, B.M., and Lipinski, J.F. (1982): Failure to find evidence of schizophrenia in first degreee relatives of schizophrenic probands. *Am.J.Psychiatry*, 139:826-827.

Porteous, D.J., Morten, J.E.N., Cranston, G., Fletcher, J.M., Mitchell, A., van Heyningen, V., Fantes, J.A., Boyd, P.A., and Hastie, N.D. (1986): Molecular and physical arrangements of human DNA in *HRAS1*- selected, chromosome-mediated transfectants. *Mol.Cell.Biol.*, 6:2223-2232.

Prescott, C.A. and Gottesman, I.I. (1993): Genetically mediated vulnerability to schizophrenia. *Psychiatr.Clin.North Am.*, 16:245-267.

Pulver, A.E., Liang, K.Y., Brown, C.H., Wolyniec, P., McGrath, J., Adler, L., Tam, D., Carpenter, W.T., and Childs, B. (1992): Risk factors in schizophrenia. Season of birth, gender, and familial risk. *Br.J.Psychiatry*, 160:65-71.

Pulver, A.E., Karayiorgou, M., Lasseter, V.K., Wolyniec, P., Kasch, L., Antonarakis, S., Housman, D., Kazazian, H.H., Meyers, D., Nestadt, G., Ott, J., Liang, K.Y., Lamacz, M., Thomas, M., Childs, B., Diehl, S.R., Wang, S.B., Murphy, B., Sun, C.E., Oneill, F.A., Nie, L., Sham, P., Burke, J., Duke, B.W., Duke, F., Kipps, B.R., Bray, J., Hunt, W., Shinkwin, R., Nuallain, M.N., Su, Y., Maclean, C.J., Walsh, D., Kendler, K.S., Gill, M., Vallada, H., Mant, R., Asherson, P., Collier, D., Parfitt, E., Roberts, E., Nanko, S., Walsh, C., Daniels, J., Murray, R., McGuffin, P., Owen, M., Laurent, C., Dumas, J.B., Damato, T., Jay, M., Martinez, M., Campion, D., and Mallet, J. (1994): Follow-up of a report of a potential linkage for schizophrenia on chromosome 22q12-q13.1 .2. *American Journal Of Medical Genetics*, 54:44-50.

Pulver, A.E., Lasseter, V.K., Kasch, L., Wolyniec, P., Nestadt, G., Blouin, J.L., Kimberland, M., Babb, R., Vourlis, S., Chen, H.M., Lalioti, M., Morris, M.A., Karayiorgou, M., Ott, J., Meyers, D., Antonarakis, S.E., Housman, D., and

Kazazian, H.H. (1995): Schizophrenia - a genome scan targets chromosomes 3p and 8p as potential sites of susceptibility genes. *American Journal Of Medical Genetics*, 60:252-260.

Quackenbush, J., Davies, J., Bailis, J.M., Khristich, J.V., Diggle, K., Marchuck, Y., Tobin, J., Clark, S.P., Rodkins, A., Marcano, S., Chirukian, A.C., Hutchinson, J.S., Probst, S., Romberg, L., Wei, Y.H., Nowak, N.J., Garner, H.R., Smith, M.W., Selleri, L., and Evans, G.A. (1995): An STS content map of human chromosome 11: localization of 910 YAC clones and 109 islands. *Genomics*, 29:512-525.

Rao, M.L. and Moller, H.J. (1994): Biochemical findings of negative symptoms in schizophrenia and their putative relevance to pharmacologic treatment. A review. *Neuropsychobiology*, 30:160-172.

Reynolds, G.P., Mason, S.L., Meldrum, A., Dekeczer, S., Parnes, H., Eglen, R.M., and Wong, E.H.F. (1995): 5-hydroxytryptamine (5-ht)(4) receptors in post-mortem human brain- tissue - distribution, pharmacology and effects of neurodegenerative diseases. *British Journal Of Pharmacology*, 114:993-998.

Riecher Rossler, A., Hafner, H., Dutsch Strobel, A., Oster, M., Stumbaum, M., van Gulick Bailer, M., and Loffler, W. (1994): Further evidence for a specific role of estradiol in schizophrenia? *Biol.Psychiatry*, 36:492-494.

Rifkin, L., Lewis, S., Jones, P., Toone, B., and Murray, R. (1994): Low birth weight and schizophrenia. *Br.J.Psychiatry*, 165:357-362.

Ritsner, M., Sherina, O., and Ginath, Y. (1992): Genetic epidemiologic-study of schizophrenia - reproduction behavior. *Acta Psychiatrica Scandinavica*, 85:423-429.

Roberts E Neurosci Res Prog Bull 1972 10 468-81

Rommens, J.M., Iannuzzi, M.C., Kerem, B.-S., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., and Hidaka, N. (1989): Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*, 245:1059-1065.

Rosier, M.-F., Reguigne-Arnould, I., Couillin, P., Devignes, M.-D., and Auffray, C. (1995): Regional Assignment of 68 new human gene transcripts on chromosome 11. *Genome Research*, 5:60-70.

Roth, B.L. (1994): Multiple serotonin receptors: clinical and experimental aspects. *Ann.Clin.Psychiatry*, 6:67-78.

Rubinstein, M., Phillips, T.J., Bunzow, J.R., Falzone, T.L., Dziewczapolski, G., Zhang, G., Fang, Y., Larson, J.L., McDougall, J.A., Chester, J.A., Saez, C., Pugsley, T.A., Gershanik, O., Low, M.J., and Grandy, D.K. (1997): Mice lacking dopamine D4 receptors are supersensitive to ethanol, cocaine, and methamphetamine. *Cell*, 90:991-1001.

Rubinstein, P., Walker, M.E., and Ginsbergfellner, F. (1981): The hla system as genetic-marker for insulin-dependent diabetes. *Archivos De Biologia Y Medicina Experimentales*, 14:33-34.

Rubinsztein, D.C., Leggo, J., Goodburn, S., Crow, T.J., Lofthouse, R., DeLisi, L.E., Barton, D.E., and Ferguson Smith, M.A. (1994): Study of the Huntington's disease (HD) gene CAG repeats in schizophrenic patients shows overlap of the normal and HD affected ranges but absence of correlation with schizophrenia. *J.Med.Genet.*, 31:690-693.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988): Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239:487-491.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning. A laboratory manual.* CSH Laboratory Press, Cold Spring Harbor.

Sanders, A.R., Hamilton, J.D., Fann, W.E., and Patel, P.I. (1991): Association of gentic variation in the porphobilinogen deaminase gene with schizophrenia. *Am.J.Hum.Genet.*, 49:A2011-A2011

Sartorius, N., Jablensky, A., Korten, A., Ernberg, G., Anker, M., Cooper, J.E., and Day, R. (1986): Early manifestations and first-contact incidence of schizophrenia in different cultures. A preliminary report on the initial evaluation phase of the WHO Collaborative Study on determinants of outcome of severe mental disorders. *Psychol.Med.*, 16:909-928.

Schuler, G.D., Boguski, M.S., Stewart, E.A., Stein, L.D., Gyapay, G., Rice, K., White, R.E., Rodrigueztome, P., Aggarwal, A., Bajorek, E., Bentolila, S., Birren, B.B., Butler, A., Castle, A.B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P.J.R., Dibling, T., Drouot, N., Dunham, I., Duprat, S., East, C., Edwards, C., Fan, J.B., Fang, N., Fizames, C., Garrett, C., Green, L., Hadley, D., Harris, M., Harrison, P., Brady, S., Hicks, A., Holloway, E., Hui, L., Hussain, S., Louisditsully, C., Ma, J., Macgilvery, A., Mader, C., Maratukulam, A., Matise, T.C., Mckusick, K.B., Morissette, J., Mungall, A., Muselet, D., Nusbaum, H.C., Page, D.C., Peck, A., Perkins, S., Piercy, M., Qin, F., Quackenbush, J., Ranby, S., Reif, T., Rozen, S., Sanders, C., She, X., Silva, J., Slonim, D.K., Soderlund, C., Sun, W.L., Tabar, P., Thangarajah, T., Vegaczamy, N., Vollrath, D., Voyticky, S., Wilmer, T., Wu, X., Adams, M.D., Auffray, C., Walter, N.A.R., Brandon, R., Dehejia, A., Goodfellow, P.N., Houlgatte, R., Hudson, J.R., Ide, S.E., Iorio, K.R., Lee, W.Y., Seki, N., Nagase, T., Ishikawa, K., Nomura, N., Phillips, C., Polymeropoulos, M.H., Sandusky, M., Schmitt, K., Berry, R., Swanson, K., Torres, R., Venter, J.C., Sikela, J.M., Beckmann, J.S., Weissenbach, J., Myers, R.M., Cox, D.R., James, M.R., Bentley, D., Deloukas, P., Lander, E.S., and Hudson, T.J. (1996): A gene map of the human genome. *Science*, 274:540-546.

Seeman, P., Ulpian, C., Chouinard, G., Van Tol, H.H., Dwosh, H., Lieberman, J.A., Siminovitch, K., Liu, I.S., Waye, J., Voruganti, P., Hudson, C., Serjeant, G.R., Masibay, A.S., and Seeman, M.V. (1994): Dopamine D4 receptor variant, D4GLYCINE194, in Africans, but not in Caucasians: no association with schizophrenia. *Am.J.Med.Genet.*, 54:384-390.

Seeman, P. and Vantol, H.M. (1994): Dopamine-receptor pharmacology. *Trends In Pharmacological Sciences*, 15:264-270.

Shaikh, S., Gill, M., Owen, M., Asherson, P., McGuffin, P., Nanko, S., Murray, R.M., and Collier, D.A. (1994): Failure to find linkage between a functional polymorphism in the dopamine D4 receptor gene and schizophrenia. *Am.J.Med.Genet.*, 54:8-11.

Sham, P.C. and Curtis, D. (1995): Monte-carlo tests for associations between disease and alleles at highly polymorphic loci. *Annals Of Human Genetics*, 59:97-105.

Sherrington, R., Brynjolfson, J., Peturson, H., Potter, M., Dudleston, K., Barraclough, B., Wasmuth, J., Dobbs, M., and Gurling, H. (1988): Localisation of a susceptibility locus for schizophrenia on chromosome 5. *Nature.*, 336:164-167.

Sikela, J.M. and Auffray, C. (1993): Finding new genes faster than ever. *Nature Genetics*, 3:189-191.

Smith, M., Wasmuth, J., McPherson, J.D., Wagner, C., Grandy, D., Civelli, O., Potkin, S., and Litt, M. (1989): Cosegregation of an 11q22-9p22 translocation with affective disorder: proximity of the dopamine D2 receptor gene relative to the translocation breakpoint. *Am.J.Hum.Genet.*, 45:A220-A220

Smith, GN., Honer, WG., Kopala, L., MacEwan, GW., Altman, S., Smith, A. (1995). Obstetric complications and severity of illness in schizophrenia. *Schizophrenia Research.* 14(2):113-120.

Sobue, K. (1993): Actin-based cytoskeleton in growth cone activity. *Neurosci.Res.*, 18:91-102.

Sommer, S.S., Lind, T.J., Heston, L.L., and Sobell, J.L. (1993): Dopamine D4 receptor variants in unrelated schizophrenic cases and controls. *Am.J.Med.Genet.*, 48:90-93.

Southern, E.M. (1975): Detection of specific sequences among DNA fragments separated by gel elctrophoresis. *J.Mol.Biol.*, 137:120-124.

Southern, E.M. (1996): Dna chips - analyzing sequence by hybridization to oligonucleotides on a large-scale. *Trends in Genetics*, 12:110-115.

Spitzer, R.L., Endicott, J., and Robins, E. (1978): *Research diagnostic criteria for a selected group of functional disorders*. New York State Psychiatric Institute, New York.

Spring, B. (1981): Stress and schizophrenia - some definitional issues. *Schizophrenia Bulletin*, 7:24-33.

St.Clair, D., Blackwood, D., Muir, W., Carothers, A., Walker, M., Spowart, G., Gosden, C., and Evans, H.J. (1990): Association within a family of a balanced autosomal translocation with major mental-illness. *Lancet*, 336:13-16.

St.Clair, D. (1994): Expanded CAG trinucleotide repeat of Huntington's disease gene in a patient with schizophrenia and normal striatal histology [letter]. *J.Med.Genet.*, 31:658-659.

St.Clair, D., Muir, W., and Blackwood, D. (1994): Molecular Biology of Schizophrenia. In: *Molecular and Cell Biology of Neuropsychiatric Diseases*, edited by F. Owen, et al, pp. 160-172. Chapman and Hall, London.

Staden, R. (1996): The staden sequence-analysis package. *Molecular Biotechnology*, 5:233-241.

Stclair, D. (1994): Expanded cag trinucleotide repeat of huntingtons-disease gene in a patient with schizophrenia and normal striatal histology. *Journal Of Medical Genetics*, 31:658-659.

Stoffelmayr, B.E. and Hunter, J.E. (1983): Premorbid functioning and outcome in schizophrenia: a cumulative analysis. *Journal of Consulting and Clinical Psychology*, S1:338-352.

Strange, P.G. (1992): Schizophrenia. In: *Brain Biochemistry and Brain Disorders*, Anonymouspp. 226-257. Oxford University Press, New York.

Straub, R.E. (1995): A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. *Nature Genetics*, 11:287-293.

Stromgren, E. (1987): Changes in the incidence of schizophrenia? *Br.J.Psy.*, 150:1-7.

Su, L.K., Kinzler, K.W., Vogelstein, B., Preisinger, A.C., Moser, A.R., Luongo, C., Gould, K.A., and Dove, W.F. (1992): Multiple intestinal neoplasia caused by a mutation in the murine homolog of the apc gene. *Science*, 256:668-670.

Suddath, R.L., Christison, G.W., Torrey, E.F., Casanova, M.F., and Weinberger, D.R. (1990): Anatomical abnormalities in the brains of monozygotic twins discordant for schizophrenia. *New England Journal of Medicine*, 322:789-794.

Susser, E., Neugebauer, R., Hoek, H., Brown, A., Lin, S., Labowitz, D., and Gorman, J. (1995): An epidemiologic study of schizophrenia after prenatal exposure to famine. *Schizophr.Res.*, 15(1-2):197-197.

Syvalahti, E.G. (1994): Biological factors in schizophrenia - structural and functional-aspects. *British Journal of Psychiatry*, 164:9-14.

Tahvanainen, E., Beggs, A.H., and Wallgrenpettersson, C. (1994): Exclusion of 2 candidate loci for autosomal recessive nemaline myopathy. *Journal Of Medical Genetics*, 31:79-80.

Takei, N., Sham, P.C., Ocallaghan, E., and Murray, R.M. (1992): Cities, winter birth, and schizophrenia. *Lancet*, 340:558-559.

Tandon, R. and Greden, J.F. (1989): Cholinergic hyperactivity and negative schizophrenic symptoms - a model of cholinergic dopaminergic interactions in schizophrenia. *Archives Of General Psychiatry*, 46:745-753.

Taylor, G.R. (1991): Polymerase chain reaction: basic principles and automation. In: *PCR: A practical approach*, edited by M.J. McPherson, et al, pp. 1-14. IRL Press at Oxford University Press, Oxford.

Taylor, M.A. and Amir, N. (1995): Sinister psychotics - left-handedness in schizophrenia and affective- disorder. *Journal Of Nervous And Mental Disease*, 183:3-9.

Taylor, P. and Fleminger, J.J. (1981): The lateralization of symptoms in schizophrenia. *British Journal Of Medical Psychology*, 54:59-65.

Terwilliger, J.D.O.J. (1994): Linkage disequilibrium between alleles at marker loci. In: *Handbook of Human Genetic Linkage*, Anonymouspp. 188-210. The Johns Hopkins University Press, Baltimore, Maryland.

Tienari, P. (1963): Psychiatric illnesses in identical twins. *Acta Psychiatr.Scand.Suppl.*,

Tienari, P., Wynne, L.C., Moring, J., Lahti, I., Naarala, M., Sorri, A., Wahlberg, K.E., Saarento, O., Seitamaa, M., Kaleva, M., and Laksy, K. (1994): The Finnish adoptive family study of schizophrenia. Implications for family research. *Br.J.Psychiatry Suppl.*, 20-26.

Todd, J.A. (1990): The role of mhc class-ii genes in susceptibility to insulin-dependent diabetes-mellitus. *Current Topics In Microbiology And Immunology*, 164:17-40.

Todd, J.A., Aitman, T.J., Cornall, R.J., Ghosh, S., Hall, J., Hearne, C.M., Knight, A., Love, J., Mcaleer, M.A., Prins, J.B., Lathrop, M., Peterson, L., and Wicker, L. (1991): Genetic-analysis of a complex, multifactorial disease, autoimmune type-1 (Insulin-dependent) Diabetes. *Research In Immunology*, 142:483-483.

Torgersen, S. (1984): Genetic and nosological aspects of schizotypal and bordeline personality disorders: a twin study. *Arch Gen Psych*, 41:546-554.

Torgersen, S., Onstad, S., Skre, I., Edvardsen, J., and Kringlen, E. (1993): True schizotypal personality-disorder - a study of co-twins and relatives of schizophrenic probands. *American Journal of Psychiatry*, 150:1661-1667.

Toru, M., Kurumaji, A., and Ishimaru, M. (1994): Excitatory Amino Acids: Implications for Psychiatric Disorders Research. *Life Sciences*, 55:1683-1699.

Tsuang, M.T., Gilbertson, M.W., and Faraone, S.V. (1991): The genetics of schizophrenia - current knowledge and future- directions. *Schizophrenia Research*, 4:157-171.

Turner, E., Ewing, J., Shilling, P., Smith, T.L., Irwin, M., Schuckit, M., and Kelsoe, J.R. (1992): Lack of association between an rflp near the d2 dopamine receptor gene and severe alcoholism. *Biological Psychiatry*, 31:285-290.

Vallada, H.P., Gill, M., Sham, P., Lim, L.C., Nanko, S., Asherson, P., Murray, R.M., McGuffin, P., Owen, M., and Collier, D. (1995): Linkage studies on chromosome-22 in familial schizophrenia. *American Journal Of Medical Genetics*, 60:139-146.

van Tol, H.H.M., Wu, C.M., Guan, H.C., Ohara, K., Bunzow, J.R., Civelli, O., Kennedy, J., Seeman, P., Niznik, H.B., and Jovanovic, V. (1992): Mulitple dopamine D4 receptor variants in the human population. *Nature*, 358:149-152.

Velakoulis, D. and Pantelis, C. (1996): What have we learned from functional imaging studies in schizophrenia - the role of frontal, striatal and temporal areas. *Australian And New Zealand Journal Of Psychiatry*, 30:195-209.

Vincent, J.B., Klempan, T., Parikh, S.S., Sasaki, T., Meltzer, H.Y., Sirugo, G., Cola, P., Petronis, A., and Kennedy, J.L. (1996): Frequency-analysis of large cag/ctg trinucleotide repeats in schizophrenia and bipolar affective-disorder. *Molecular Psychiatry*, 1:141-148.

Waddington, J.L. (1993): Schizophrenia: developmental neuroscience and pathobiology. *Lancet*, 341:531-536.

Walker, E.F., Lewine, R.R.J., and Neumann, C. (1996): Childhood behavioral-characteristics and adult brain morphology in schizophrenia. *Schizophrenia Research*, 22:93-101.

Walker, E.F. (1996): Behavioral-studies in neurodevelopment. *Biological Psychiatry*, 39:196-196.

Wang, S., Sun, C., Walczak, C.A., Ziegle, J.S., Kipps, B.R., Goldin, L.R., and Diehl, S.R. (1995): Evidence for a susceptibility locus for schizophrenia on chromosome 6pter-p22. *Nat.Genet.*, 10:41-46.

Wang, S. (1995): Evidence for a susceptibility locus for schizophrenia on chromosome 6pter-p22. *Nature Genetics*, 10:41-46.

Wang, Z.W., Black, D., Andreasen, N.C., and Crowe, R.R. (1993): A linkage study of chromosome-11q in schizophrenia. *Archives Of General Psychiatry*, 50:212-216.

Ward, T. and Davies, K.E. (1993): The leading role of stss in genome mapping. *Human Molecular Genetics*, 2:1097-1098.

Weil, D., Blanchard, S., Kaplan, J., Guilford, P., Gibson, F., Walsh, J., Mburu, P., Varela, A., Levilliers, J., Weston, M.D., Kelley, P.M., Kimberling, W.J., Wagenaar, M., Levi-Acobas, F., Larget-Piet, D., Munnich, A., Steel, K.P., Brown, S.D.M., and Petit, C. (1995): Defective myosin VIIA gene responsible for Usher syndrome type 1B. *Nature*, 374:60-61.

Weinberger, D.R. (1987): Implications of normal brain-development for the pathogenesis of schizophrenia. *Archives Of General Psychiatry*, 44:660-669.

Williams, J., Farmer, A.E., Wessely, S., Castle, D.J., and McGuffin, P. (1993): Heterogeneity in schizophrenia - an extended replication of the hebephrenic-like and paranoid-like subtypes. *Psychiatry Research*, 49:199-210.

Williams, J., Spurlock, G., McGuffin, P., Mallet, J., Nothen, M.M., Gill, M., Aschauer, H., Nylander, P.O., Macciardi, F., and Owen, M.J. (1996): Association between schizophrenia and t102c polymorphism of the 5- hydroxytryptamine type 2a-receptor gene. *Lancet*, 347:1294-1296.

Williams, R.M., Shear, J.B., Zipfel, W.R., Maiti, S., and Webb, W.W. (1997): Three-photon excitation imaging of serotonin secretion by RBL-2H3 cells. *Biophysical Journal*, 72:TU100-TU100

Wilson, A.F., Baileywilson, J.E., Bamba, V., and Siervogel, R.M. (1993): Possible evidence of linkage for several traits related to body- composition. *American Journal Of Human Genetics*, 53:880-880.

Winship, P.R. (1989): An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucl.Acids Res.*, 17:1266-1266.

Wirth, J., Wagner, T., Meyer, J., Pfeiffer, R.A., Tietze, H.U., Schempp, W., and Scherer, G. (1996): Translocation breakpoints in 3 patients with campomelic dysplasia and autosomal sex reversal map more than 130 kb from sox9. *Human Genetics*, 97:186-193.

Wong DF, Wagner HN, Tune Le, Dannals RF, Pearlson GD, Links JM, Tamminga CA, Broussolle EP, Ravert HT, Wilson AA, Toung JKT, Malat J, Williams JA, Otuama LA, Snyder SH, Kuhar MJ, Gjedde A (1996) Positron emission tomography reveals elevated d2 dopamine receptors in drug-naive schizophrenics. *Science*, ,.234,.(4783),.1558-1563

Wright, P., Gill, M., and Murray, R.M. (1993): Schizophrenia - genetics and the maternal immune-response to viral- infection. *American Journal Of Medical Genetics*, 48:40-46.

Wyatt, R.J., Alexander, R.C., Egan, M.F., and Kirch, D.G. (1988): Schizophrenia, just the facts. What do we know, how well do we know it? *Schizophrenia Research.*, 1:3-18.

Wyszynski, M., Lin, J., Rao, A., Nigh, E., Beggs, A.H., Craig, A.M., and Sheng, M. (1997): Competitive binding of alpha-actinin and calmodulin to the NMDA receptor. *Nature*, 385:439-442.

Xu, Y., Mural, R.J., Einstein, J.R., Shah, M.B., and Uberbacher, E.C. (1996): Grail - a multiagent neural-network system for gene identification. *Proceedings Of The*

National Academy of science 84:1544-1552.

Yang, L., Li, T., Wiese, C., Lannfelt, L., Sokoloff, P., Xu, C.T., Zeng, Z., Schwartz, J.C., Liu, X., and Moises, H.W. (1993): No association between schizophrenia and homozygosity at the D3 dopamine receptor gene. *Am.J.Med.Genet.*, 48:83-86.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985): Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, 33:103-119.

Zatz, M. H. Vallada, M. S. Melo, M. R. Passos Bueno, A. H. Vieira, M. Vainzof, M. Gill, and V. Gentil. Cosegregation of schizophrenia with Becker muscular dystrophy: susceptibility locus for schizophrenia at Xp21 or an effect of the dystrophin gene in the brain? *J.Med.Genet.* 30:131-134, 1993.

Papers presented during the course of this thesis

Psychiatric Genetics 1997 volume 7 in press

An allelic association study of two polymorphic markers in close proximity to a balanced translocation t(1:11) which co-segregates with mental illness

J.C. Wilson-Annan[1], D.H.R. Blackwood[1,2], W. Muir[1,2], J.K. Millar[1], D.J. Porteous[1]

[1] MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK and [2]Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF

Correspondence to: Julie Wilson-Annan at above address. Tel: (0) 131 332 2741; Fax: (0) 131 343 2620

We report a case control association study using polymorphic markers D1S1621 and D11S931 in unrelated subjects with schizophrenia, unipolar depression and a matched control group. The two polymorphic markers were identified during the positional cloning of the translocation breakpoint t(1:11)(q43:q14.3) which cosegregates with schizophrenia and affective disorders. These markers provided an opportunity to investigate linkage disequilibrium with a postulated schizophrenia susceptibility gene close to the translocation breakpoint in random populations of schizophrenia and unipolar depression subjects compared to a normal control population.

No significant differences between allele frequencies for either of the markers in the affected populations were observed in comparison to the control group, evidence against a nearby gene of major effect in the populations studied.

Keywords: Association - schizophrenia - unipolar depression

INTRODUCTION

Schizophrenia is a serious and debilitating mental illness with a life time risk of ~1% in the human population. Family, twin and adoption studies have provided strong evidence for there being a genetic component involved in schizophrenia but the nature of this component is still unclear. Genetic linkage studies have failed to provided clarification of the gene(s) involved in schizophrenia, most likely due to the complex nature of the disorder, namely the polygenic non-Mendelian inheritance, probable aetiologic heterogeneity and difficulties with precise phenotype definition. However, recent world-wide collaborations and methodological refinements have enabled large numbers of families to be studied, thus improving the power of the technique and have enabled the detection of several candidate regions for involvement with schizophrenia, most notably chromosome 6p (Wang *et al*, 1995, Straub *et al* 1995) and 22q (Pulver *et al*, 1994; Coon *et al*, 1994). Further replication and refinement of these regions is needed. Population based association studies, in which the frequency of an allele at the marker locus in disease population is compared to that of an ethnically matched control population, provide an alternative approach to family based linkage analysis and have been successfully performed in a number of complex disorders such as late onset Alzheimer's disease with apolipoprotein E allele (Corder *et al* 1994) and myocardial infarction with angiotensin converting enzyme (Cambien *et al* 1992).

The identification of cytogenetic abnormalities associated with disease has played an important part in localising genes causing several human diseases. We have identified a balanced translocation t(1:11)(q42.2,q14.3) which cosegregates with schizophrenia and affective disorders in a large Scottish pedigree (St Clair *et al* 1990). A positional cloning strategy has been employed to elucidate the relationship between the translocation and the psychiatric diagnosis in this family, based on the hypothesis that a gene or genes involved in mental illness in this family reside at or near the translocation breakpoint on chromosome 1 and/or 11 (Fletcher *et al* 1993, Muir *et al* 1995, Evans *et al* 1995).

A polymorphic triplet repeat, D1S1621, has been identified which lies in close proximity to the translocation breakpoint on chromosome 1. A dinucleotide repeat polymorphic marker, D11S931, was also available on chromosome 11 in close proximity to the breakpoint.

These two markers provided an opportunity to investigate an association between the translocation breakpoint region and a possible schizophrenia susceptibility gene in a random population of affected subjects versus normal controls.

We therefore studied the frequencies of the alleles of D11S931 and D1S1621 in random populations of subjects with schizophrenia and unipolar depression (a diagnosis also prominent in the t(1;11) family) compared to ethnically matched control subjects in order to determine if there is a difference in allele frequencies between the affected populations and the control population.

METHODS

Subjects:

Subjects with schizophrenia subjects (n=105) and unipolar depression (n=84) were all either in-patients or out-patients at the Royal Edinburgh Hospital. All patients were interviewed by psychiatrists using the Schedule for Affective Disorder and schizophrenia-lifetime version (SADS-L) (Endicott and Spitzer 1978). Diagnoses were based on RDC and DSM IV criteria (Spitzer *et al* 1978 and American Psychiatric Association 1994). Control subjects were obtained from the general population and were excluded from the study if they had a personal or family history of psychiatric illness. All subjects originated from the south east of Scotland and gave informed consent to take part in this study.

Standard procedures were used to extract DNA from peripheral blood samples.

Subjects were genotyped for markers D11S931 and D1S1621 by polymerase chain reaction (PCR) using a fluorescent primer and electrophoresis was carried out using a 40-lane automated laser fluorescence (ALF) sequencer.

PCR:

Amplification of D1S1621 marker was achieved using primers 5'-TTTCTCACCTTTAAATGTCATCA-3' and 5'-CCAGTACGCAGATGGTCCTA-3'. D11S931 was amplified with oligonucleotide primers 5'- ATGTTGGTAGGTATTCT- 3' and 5'-GAGAAATAGTATGTGTTTGCC -3'. PCR for both markers was performed in a final volume of 25µl containing 200µM of each dNTP, 150ng primer, 1.5mM $MgCl_2$, 10mM Tris/HCL, pH 8.3, 50mM KCl, 0.01% (w/v) gelatin, 0.1% Triton X-100, 1Unit Taq DNA polymerase (Perkin Elmer, Norwalk, Connecticut).

PCR conditions were as follows for D1S1621: 30 seconds at 94°C, 1 min 15 sec at 55°C, 15 sec at 72°C (27 cycles), 30 sec at 94°C, 1 min 15 sec at 55°C, 6 min at 72°C (last cycle). For D11S931 PCR conditions were 5 min at 94°C, 30 sec at 47°C, 1 min at 72°C (1 cycle), 1min at 94°C, 30 sec at 47°C, 1 min at 72°C (30 cycles), 1min at 94°C, 30 sec at 47°C, 2 min at 72°C (last cycle).

Electrophoresis was carried out on a 40-lane automated laser fluorescence (ALF) sequencer (Pharmacia). Gels were made up with 6% Hydrolink (Long Ranger™, AT Biochemicals Ltd) in 0.6X TBE (0.06 M Tris base, 0.05M boric acid, 0.6mM EDTA)containing 7 M urea and run in the same buffer at constant power (55W) and constant temperature (50°C).

After the PCR reaction the products from both reactions (D11S931 and D1S1621), for the same patient, were run in the same lane on the ALF sequencer (since the products are of different size) with appropriate size markers and gold standards. Automated genotyping was carried out using 'ALF' manager and 'ALP' software.(He *et al* 1995, Mansfield *et al* 1994). All gel results were also checked manually.

The presence of an allelic association was sought using CLUMP (Sham *et al.* 1995).

RESULTS

Genotypes were obtained for 143 control subjects 100 schizophrenia subjects and 84 random unipolar depressed subjects with markers D1S1621 and D11S931. Allele frequencies in affecteds and controls for both markers tested are shown in Tables I and II.

No significant statistical difference in allele frequency or genotype was found for marker D1S1621 in either the schizophrenia subjects (χ^2 = 3.4 P= 0.65) or the unipolar depressed subjects (χ^2 = 5.0 P= 0.39) as compared to the control population. Marker D11S931 similarly did not show any statistically significant differences between the control population and either schizophrenia subjects (χ^2 = 1.6 P= 0.66) or unipolar depressed subjects (χ^2 = 1.8 P=0.63).

The D11S931 sample has 76% power to detect a 20% difference in allele frequency and 23% power to detect a 10% difference.

The D1S1621 sample has 68% power to a detect 20% difference in allele frequency and 19% power to detect a 10% difference (Cohen 1988).

TABLE I: Percentage Allele Frequency at marker D11S931. (Numbers in parentheses represent actual allele frequency)

	Allele sizes (bp)			
	149	159	163	165
Controls (n=143)	30 (86)	10 (29)	22 (64)	38 (107)
Schizophrenics (n=100)	28.5 (57)	13.5 (27)	23.5 (47)	34.5 (69)
Unipolar depression (n=84)	30 (50)	12.5 (21)	25.5 (43)	32 (54)

TABLE II: Percentage Allele frequency at marker D1S1621. (Numbers in parentheses represent actual allele frequency)

	Allele size (bp)					
	252	255	258	261	264	267
Controls (n=143)	13 (36)	25 (73)	33 (94)	16 (45)	11 (31)	2 (7)
Schizophrenia (n=105)	10 (21)	27 (57)	31 (66)	18 (37)	9 (19)	5 (10)
Unipolar Depression (n=84)	12 (20)	31.5 (53)	31.5 (53)	11 (18)	9 (16)	5 (8)

DISCUSSION

We observed no significant differences in the allele frequencies of markers D1S1621 or D11S931 in our population of schizophrenic subjects and unipolar depressed subjects in comparison to a control population.

This lack of association may indicate that the two markers investigated are not in linkage disequilibrium with a major gene of high prevalence in the population tested. Alternatively, it may be that this postulated gene in close proximity to the translocation breakpoint is of major effect in the translocation family but is prominent in only a minority of the affected population studied and is therefore not of strong enough effect to be picked up in the association study.

It is unlikely that we have failed to demonstrate an association due to population stratification as the control and affected populations have been selected from the same geographical region as far as possible. Family based controls should however be considered wherever possible to lessen this problem (Rubinstein *et al* 1981, Falk and Rubinstein 1987).

In conclusion, the present study does not reveal an association between the markers D1S1621 and D11S931 with either schizophrenia or unipolar depression in this Scottish population, evidence against a nearby gene of major effect.

REFERENCES

1. American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders*, 4th edition, American Psychiatric Association, Washington D.C
2. Cambien F, Poirier O, Lecerf L, Evans A, Cambou J, Arveiler D, Luc G, Bard J,Bara L, Ricard S, Tiret L, Amouyel P, Alhenc-Gelas F and Soubrier F (1992) Deletion

polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature,* **359**, 641-644

3. Coon H, Holik J, Hoff M, Reimherr F, Wender P, Myles Worsley M, Waldo M, Freedman R, Byerley W (1994) Analysis of chromosome 22 markers in nine schizophrenia pedigrees. *Am J Med Genet,* **54**, 72-79
4. Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmachel DE, Gaskell PC, Rimmler JB, Lacke PA, Conneally PM, Schmader KE, Small GW, Roses AD, Haines JL and Pericak-Vance MA (1994) Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nature Genetics,***7**, 180-184
5. Endicott J and Spitzer R (1978) A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch General Psychiatry,* **35**, 837-844
6. Evans KL, Brown J, Shibasaki Y, Devon RS, He L, Arveiler B, Christie S, Maule JC, Baillie D, Slorach EM, Anderson SM, Gosden JR, Petit J, Weith A, Gosden CM, Blackwood DHR, St. Clair D, Muir W, Brookes AJ, Porteous DJ (1995) A continuous clone map over 3Mb on the long arm of chromosome 11 across a balanced translocation associated with schizophrenia. *Genomics,* **28**, 420-428
7. Falk CT, Rubinstein P (1987) Haplotype Relative Risk: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet,* **51**, 227-233
8. Fletcher JM, Evans KL, Baillie D, Byrd P, Hanratty D, Leach S, Julier C, Gosden JR, Muir W, Porteous DJ, St. Clair D, van Heyningen V (1993) Schizophrenia-associated chromosome 11q21 translocation: Identification of flanking markers and development of chromosome 11q fragment hybrids as cloning and mapping resources. *Am J Hum Genet,* **52**, 478-490
9. He L, Mansfield DC, Brown AF, Green DK, Morris SW, St Clair DM, Muir WJ, Maclean A, Wright AF, Blackwood DHR (1995) Automated linkage analysis in psychiatric disorders. *Am J Med Genet,* **60**, 192-198

10. Mansfield DC, Brown AF, Green DK, Carothers AD, Morris SW, Evans HJ and Wright AF (1994) Automation of genetic linkage analysis using fluorescent microsatellite markers. *Genomics*, **24**, 225-233

11. Muir W, Gosden CM, Brookes AJ, Fantes J, Evans KL, Maguire SM, Stevenson B, Boyle S, Blackwood DH, St.Clair DM, Porteous DJ, Weith A (1995) Direct microdissection and microcloning of a translocation breakpoint region, t(1;11)(q42.2;q21), associated with schizophrenia. *Cytogenet Cell Genet,***70**, 35-40

12. Pulver AE, Karayiorgou M, Wolyneic PS, Lasseter VK, Kasch L, Nestadt G, Antonarakis S, Housman D, Kazazian HH, Meyers D, Ott J, Lamacz M, Liang K, Hanfelt J, Ullrich G, Dimarchi N, Ramu E, McHugh PM, Adler L, Thomas M, Carpenter WT, Mansschreck T, Gordon CT, Kimberland M, Babb, Puck J, Childs B (1994) Sequential strategy to identify a susceptibility gene for schizophrenia: Report of potential linkage on chromosome 22q12-q13: Part 1. *Neuropsych. Genet,* **54**, 36-43

13. Rubinstein P, Walker M, Carpenter C, Carrier C Krassner J. Falk C, Ginsberg F (1981) Genetics of HLA disease association: The use of the haplotype relative risk (HRR) and the "Haplo-Delta" (Dh) estimates in juvenile diabetes from three racial groups. *Hum Immunol,* **3**,384

14. Sham PC and Curtis D (1995) Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet*, **59**, 97-105

15. Spitzer RL, Endicott J and Robins E (1978) *Research Diagnostic Criteria for a selected group of functional disorders*, 3rd edn. New York State Psychiatric Institute, New York.

16. St. Clair D, Blackwood D, Muir W, Carothers A, Walker M, Spowart G, Gosden C, Evans HJ (1990) Association within a family of a balanced autosomal translocation with major mental illness. *Lancet*, **336**, 13-16

17. Straub RE (1995) A potential vulnerability locus for schizophrenia on chromosome 6p24-22 - evidence for genetic-heterogeneity. *Nature Genetics*, **11,** 287-293

18. Wang S, Sun C, Walczak CA, Ziegle JS, Kipps BR Goldin LR, Diehl SR (1995) Evidence for a susceptibility locus for schizophrenia on chromosome 6pter-p22. *Nature Genetics*, **10**, 41-46